

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2004, 02:02:11 ; Search time 31 Seconds
(without alignments)
622.842 Million cell updates/sec

Title: US-09-926-805-7

Perfect score: 1949

Sequence: 1 GSMHSHRTIDARSEQDLL.....KGLEASTPSVLSQPLQLFLD 374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*

2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*

3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*

4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*

5: /cgn2_6/ptodata/2/iaa/PCFUS COMB.pep:*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1380	70.8	459	1	US-08-087-732-2
2	1380	70.8	459	6	Sequence 2, Appli Patent No. 5210189
3	1361	69.8	368	3	US-09-000-092-4
4	1360	69.8	368	3	US-09-000-092-6
5	1358	69.7	368	3	US-09-000-092-10
6	1316	67.5	368	3	US-09-000-092-2
7	1305	67.0	368	3	US-09-000-092-8
8	536	27.5	945	4	US-03-198-452A-1030
9	103.5	5.3	336	4	US-09-134-001C-5549
10	101.5	5.2	2318	3	US-09-091-219-24
11	101.5	5.2	2318	4	US-09-660-541-24
12	97	5.0	426	4	US-09-134-000C-4401
13	96.5	5.0	575	4	US-09-328-352-7721
14	95	4.9	1529	4	US-03-342-283C-396
15	93.5	4.8	453	4	US-09-252-991A-31268
16	92	4.7	1525	3	US-09-191-647-2
17	92	4.7	1525	3	US-09-540-245A-2
18	92	4.7	1525	3	US-09-540-153-2
19	91.5	4.7	684	4	US-09-564-805-233
20	88.5	4.5	419	4	US-09-252-991A-17834
21	88	4.5	608	4	US-09-134-001C-3395
22	88	4.5	705	4	US-09-328-352-7436
23	87.5	4.5	910	4	US-08-997-685A-2
24	87.5	4.5	1032	4	US-09-976-594-214
25	87.5	4.5	2254	2	US-08-677-010-3
26	87.5	4.5	2254	2	US-08-790-519-3
27	87	4.5	389	4	US-09-107-532A-6456

Sequence 2, Appli
Sequence 2, Appli
Sequence 11, Appli
Sequence 62, Appli
Sequence 2, Appli
Sequence 16, Appli
Sequence 3693, Ap
Sequence 27566, A
Sequence 12, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 54, Appli
Sequence 53, Appli
Sequence 5110, Ap
Sequence 6678, Ap
Sequence 20, Appli

ALIGNMENTS

RESULT 1

US-08-087-732-2

; Sequence 2, Application US/08087732

; Patent No. 5516667

; GENERAL INFORMATION:

; APPLICANT: Nishizawa, Osamu

; TITLE OF INVENTION: CHILLING RESISTANT PLANTS AND THEIR

; TITLE OF INVENTION: PRODUCTION

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/087,732

; FILING DATE: 16-JUL-1993

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP92/00024

; FILING DATE: 14-JAN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 15883/1991

; FILING DATE: 16-JAN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 283807/1991

; FILING DATE: 04-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 459 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-087-732-2

Query Match 70.8%; Score 1380; DB 1; Length 459;
Best Local Similarity 69.2%; Pred. No. 2.8e-140;

Matches	252;	Conservative	57;	Mismatches	55;	Indels	0;	Gaps	0;
QY	6	HSRTFIDARSEQDLLSGIQIRELEAGTLLPKHIAQAMEELYONYKNAVILQSAAPHAEDIVLS	65						
Db	94	HSRTFLDARSEQDLLSGIQIKEAEAGRLPANVAAAGMEELYNNYKNAVILSSGASRADETUVS	153						
QY	66	NMVAEDRMFLDVKPEPEFSPYHEALLEPNNYMFQGYIRPLVNPRESYGVNVFQVM	125						
Db	154	NMVAEDRMFLGVGDEPTTFNPYHKAVERPDYFNVTYIRPLDFFKNYVGNASIFSEL	213						
QY	126	BEOLKQDQKVLVLSNHQTEADPAVIALMLETNPHISENIYVAGDRVITDPLCKPFSMG	185						
Db	214	EDKIRQGHNVILSNHQSEADPAVISLLEAQSPFIGNIKCVAGDRVITDPLCKPFSMG	273						
QY	186	RNLLCVYKXKHMNDVPELAEMKPSNTRSLKEMALLRGSKILTIWAPSGGRORPDITN	245						
Db	274	RNLLCVYKXKHMNDVPELVDMKRNKTRSLKEMATMLRSGQLTWIAPSGGRORPNPSTG	333						
QY	246	QWPPAPFDATSLDNMRRLVDHAGLVGHYIPLAIIICHDIMPPPLQVEKEIGEKKLISPHGT	305						
Db	334	EWPPAPFDASSVDMNRRLVEHSGAPGHIYPMSSLICYDIMPPPPQVEKEIGEKKLVGHGT	393						
QY	306	GISVAPENIQEVTASCSPBEAKAANYQALYDSVCQYKVLHSAVHGKGLEASTSVS	365						
Db	394	GLSIAPENISDVTADCESPNEAKAYSQALYSVNEQYEILNSAIKRRGVEASTSRVS	453						
QY	366	LSQP	369						
Db	454	LSQP	457						
RESULT 2									
5210189-2									
; Patent No. 5210189									
; APPLICANT: MURATA, NORIO									
; TITLE OF INVENTION: DNA SEQUENCE ENCODING GLYCEROL									
; 3-PHOSPHATE ACYLTRANSFERASE									
; NUMBER OF SEQUENCES: 4									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/07/587,676									
; FILING DATE: 25-SEP-1990									
; SEQ ID NO:2:									
; LENGTH: 459									
5210189-2									

Query Match	69.8%;	Score 1361;	DB 3;	Length 368;
Best Local Similarity	58.1%;	Pred. No. 2.2e-138;		
Matches 250;	Conservative 64;	Mismatches 53;	Indels 0;	Gaps 0;
Qy	3	MHGSRTFIDARSEODLLSGIQRELEAGTLPKHIAQAAMEELYQNYKNVAALQCSAAAPHREDI	62	
Db	1	MASHSRKFLDVRBEEBELLSCIKKETEACKLPNNVAGMEELYQYRNAVIESGNPKADEI	60	
Qy	63	VLSNMVRVAFDRMFLDVKEPFEPFYPYHAILLEPNFYMFQNYIIRPLNNFRESYGVNVSVP	122	
Db	61	VLSNMTVALDRILLVDVEDPFVFSHSHKAIREFDYIFGQNYIRPLDIFGNSFVGNLSLF	120	
Qy	123	GVMEEOLKQDGKVVVLSLNHOTEADPAVIALMLETNNPHINENIITYAGDVITDPLCKPF	182	
Db	121	KDITEKLQGDNIILMSHQSEADPAVIALLEKNTSLIAENLIYIAGDVITDPLCKPF	180	
Qy	183	SMGRNLLCVYSKGMNDVPVLAEMWKGSRNTRSIKEMALLLRGGSKIITWAPSGGRDRDP	242	
Db	181	SMGRNLLCVYSKGMHYDDPDLVDVVKRANTRSLKELVLLLRGGSKIITWAPSGGRDPDA	240	
Qy	243	ITNQWFPAPDPATSLONMRRLVDHAGLVGHYIPLAILCHDIMPPLQVEKEIGEKRLISF	302	

Db 241 VTGEWYPAFDASSVDNMRLLIQHS DVPGLFPLALLCHDIMPSPQVEIEGKRVIAF 300
QY 303 HGTGISVAPEINQEVTVASCGSPPEAKAAYSQALYDSVCQYKVLHSAVHGKGLRSTP 362
Db 301 NGAGLSVAPEISPEETAAATHKNPEEVREAYSKALFDSVAMQYNVLTATISGKQGLGASTA 360
QY 363 SVLSQP 369
Db 361 DVLSQP 367

RESULT 4
US-09-000-092-6
; Sequence 6, Application US/09000092
; Patent No. 6160203
; GENERAL INFORMATION:
; APPLICANT: FERRI, Stefano
; APPLICANT: TOGURI, Toshihiro
; TITLE OF INVENTION: DNA STRANDS CODING FOR
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,092
; FILING DATE: 26-JAN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; FILING DATE: 03-JUL-1996
; APPLICATION NUMBER: JP 192123/1995
; FILING DATE: 27-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-000-092-6

Query Match 69.8%; Score 1360; DB 3; Length 368;
Best Local Similarity 68.1%; Pred. No. 2.8e-138;
Matches 250; Conservative 64; Mismatches 53; Indels 0; Gaps 0;

QY 3 MHGHSRTFIDARSEQDLSGIQRELEAGTLPKHIAQAMEELYQYKNAVLSQAAPHAEDI 62
Db 1 MASHSRKFLDVRSEELLSCKKETEAGKLPNNVAGMEELYQYRNVAIESGNPKADEI 60

QY 63 VLSNMRVAFRMLFDVKEPEFSPYHEAILEPNNYMFQNYIRPLVNFRESYVGNVSF 122
Db 61 VLSNMTVALDRILLDVEDPVSFSSHKAIREPEFYIFQNYIRPLIDFGNSFVGNLSLF 120

QY 123 GVMEEQLKQGDQVLLSNHQTEADPAVIALMLETTPHISENLIYVAGDRVITDPLCKPF 182

Db 121 KOLEEKLKQGDNIILMSNHQSEADPAVIALLEKTNLSIAENLIYTAGDRVITDPLCKPF 180
QY 183 SMGRNLLCVYSKGMNDVPELAEMKCRNTRSLKEMALLLRGSKTIWIAPSGGRDRPDP 242
Db 181 SMGRNLLCVYSKGMNDVPELAEMKCRNTRSLKEMALLLRGSKTIWIAPSGGRDRPDP 240
QY 243 ITNWFPAFPDATSLDNMRRLVDHAGLVGHYPLATLCHDIMPPLQVKEIEGKRLISF 302
Db 241 STGEWYPAFDASSVDNMRLLIQHS DVPGLFPLALLCHDIMPSPQVEIEGKRVIAF 300
QY 303 HGTGISVAPEINQEVTVASCGSPPEAKAAYSQALYDSVCQYKVLHSAVHGKGLRSTP 362
Db 301 NGAGLSVAPEISPEETAAATHKNPEEVREAYSKALFDSVAMQYNVLTATISGKQGLGASTA 360
QY 363 SVLSQP 369
Db 361 DVLSQP 367

RESULT 5
US-09-000-092-10
; Sequence 10, Application US/09000092
; Patent No. 6160203
; GENERAL INFORMATION:
; APPLICANT: FERRI, Stefano
; APPLICANT: TOGURI, Toshihiro
; TITLE OF INVENTION: DNA STRANDS CODING FOR
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,092
; FILING DATE: 26-JAN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP96/01844
; FILING DATE: 03-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 192123/1995
; FILING DATE: 27-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-000-092-10

Query Match 69.7%; Score 1358; DB 3; Length 368;
Best Local Similarity 68.4%; Pred. No. 4.6e-138;
Matches 249; Conservative 64; Mismatches 51; Indels 0; Gaps 0;

QY 6 HSRFTFDARSEQDLSGIQRELEAGTLPKHIAQAMEELYQYKNAVLSQAAPHAEDI 65
Db 4 HSRKFLDVRSEELLSCKKETEAGKLPNNVAGMEELYQYRNVAIESGNPKADEI 63

QY 66 NMRVAFDRMFLDVKEPPEFSPYHEAILEPENNYMFGONYIRPLVNPRESYVGNVSVFGVM 125
 Db 64 NMTVALDRILLDVEDPFFVSSHKAIREPPDYIFGQYIRPLDFGNSVFNLSLFKDI 123
 QY 126 BEQKQGDQKVVILSNHQTADPAVIALMLETNPHISENIYVAGDRVITDPLCKPFSMG 185
 Db 124 EEKQKQGDNIILMSNHQSEADPAVIALLEKTNLSIAENLIYIAGDRVITDPLCKPFSMG 183
 QY 186 RNLCLVYSKKHMNDVPELAEMKKESNTRSLKEMALLRGSKTIWIAPSGGRDRPDITN 245
 Db 184 RNLCLVYSKKHMYDDPELVDDVKKRANTRSLKELVLLRGSKTIWIAPSGGRDRPDVATG 243
 QY 246 QWFPAPDATSLDNMRRLVDHAGLVGHYIPLAILCHDIMPPLQVEKEIGEKLISPHGT 305
 Db 244 EWPAPDASSVDNMRLIQHSDVPGLHPLPALLCHDIMPSPQVEIEGEKVIAPNGA 303
 QY 306 GISVAPINQEVTVASCGSPPEEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTPSVS 365
 Db 304 GLSVAPISFEETIAATHKNPEEVREAYSKALFDSVAMQYNVLKTAISGKGLGASTADVS 363
 QY 366 LSQP 369
 Db 364 LSQP 367

RESULT 6

US-09-000-092-2

; Sequence 2, Application US/09000092

; Patent No. 6160203

; GENERAL INFORMATION:

; APPLICANT: FERRI, Stefano

; APPLICANT: TOGURI, Toshihiro

; TITLE OF INVENTION: DNA STRANDS CODING FOR

; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FOLEY & LARDNER

; STREET: 3000 K Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/000,092

; FILING DATE: 26-JAN-1998

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/JP96/01844

; FILING DATE: 03-JUL-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 192123/1995

; FILING DATE: 27-JUL-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Bent, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 16887/916

; TELEPHONE: (202) 672-5300

; TELEFAX: (202) 672-5399

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 368 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-000-092-2

Query Match 67.5%; Score 1316; DB 3; Length 368;
 Best Local Similarity 56.2%; Pred. No. 1.6e-133;
 Matches 243; Conservative 62; Mismatches 62; Indels 0; Gaps 0;
 QY 3 MEHGSRTFIDARSEQDLSGIQRELEAGTLPKHIAQAMBELYQYKNVAVIQAAPHAEDI 62
 Db 1 MASHSRTYRVSABEELISEIKRESIGRLPKSVAYAMEGLFHYRYNAVLSGISHADEI 60
 QY 63 VLSNMRVAFDRMFLDVKEPPEFSPYHEAILEPENNYMFGONYIRPLVNPRESYVGNVSVF 122
 Db 61 VLSNMSVMLDIEDPFPFPHKAIAREPADYISFGQYIRPLVDFGNSVGNIAIF 120
 QY 123 GVMEEQLKQGDQKVVILSNHQTADPAVIALMLETNPHISENIYVAGDRVITDPLCKPF 182
 Db 121 QMEEKQKQGDNIILMSNHQSEADPAVIALLEKTNLSIAENLIYIAGDRVITDPLCKPF 180
 QY 183 SMGRNLLCVYSKKHMNDVPELAEMKKESNTRSLKEMALLRGSKTIWIAPSGGRDRPD 242
 Db 181 SMGRNLLCVYSKKHMYDDPELVDDVKKRANTRSLKELVLLRGSKTIWIAPSGGRDRPDA 240
 QY 243 ITNOMEPAPEINQEVTVASCGSPPEEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTP 302
 Db 241 VTGEWYPAPEINQEVTVASCGSPPEEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTP 300
 QY 303 HGTGISVAPINQEVTVASCGSPPEEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTP 362
 Db 301 NGAGLSVAPISFEETIAATHKNPEEVREAYSKALFDSVAMQYNVLKTAISGKGLGASTA 360
 QY 363 SVLSQP 369
 Db 361 DVLSQP 367

RESULT 7

US-09-000-092-8

; Sequence 8, Application US/09000092

; Patent No. 6160203

; GENERAL INFORMATION:

; APPLICANT: FERRI, Stefano

; APPLICANT: TOGURI, Toshihiro

; TITLE OF INVENTION: DNA STRANDS CODING FOR

; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FOLEY & LARDNER

; STREET: 3000 K Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/000,092

; FILING DATE: 26-JAN-1998

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/JP96/01844

; FILING DATE: 03-JUL-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 192123/1995

; FILING DATE: 27-JUL-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Bent, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 16887/916

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 672-5300

; TELEFAX: (202) 672-5399

; INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-000-092-8

Query Match 67.0%; Score 1305; DB 3; Length 368;
Best Local Similarity 66.4%; Pred. No. 2.5e-132; Mismatches 60; Indels 0; Gaps 0;
Matches 241; Conservative 62;

QY 7 SRTFIDARSODLLSGTQRELEAGTLPKHIAQAMEELYQYKNAVLQSAAPHAEIVLSN 66
DB 5 SRTYRNVRSABELISEIKRESEIGRLPKSVAYAMEGLFYHRYNAVLSSGISHADEIVLSN 64
QY 67 MRVAFDRMELDVKPPFSPVHEALLPFNYYMFGQYIRPLNPNRESYVGNVSFVQME 126
DB 65 MSVWLDLFVLDIEDPFPFPHKAIKREPADYISFGQYIRPLVDFGNSYVGNIAIFOEME 124
QY 127 EQLKQGDVVLISNHOEADPAVIALMLETNPHISNIIVAGDRVITDPLCKPFSMGR 186
DB 125 EKLKQGNIIIMSNHQSADPAVIALMLETNPHISNIIVAGDRVITDPLCKPFSMGR 184
QY 187 NLLCVYSKGMNDVPELAEMKKSNTSLKEMALLRGSKIIWIAPSGGRDRPDPTNQ 246
DB 185 NLLCVYSKGMNDVPELVKXKRANTRSLKELVLLRGSKIIWIAPSGGRDRPDVTCG 244
QY 247 WFPAPDTSLDNMRRLVDHAGLVGHYIPLAILCHDIMPPLQVEKEIGKRLISHFGTG 306
DB 245 WYFAPDASSVDNRRLLIQSDVPHGLFPLALICHDIMPFPQVEIGKRVIAFNAG 304
QY 307 ISVAPEINFOEVTASCGSPPEAKAAYSOALYDSVCEQYKVLHSAVHGKGLKLEASTPSVL 366
DB 305 LSVAPEISFEIAATHKNPEVREAYSKALFDSVAMQYVNLKTAISKQGLGASTADVSL 364
QY 367 SQP 369
DB 365 SQP 367

RESULT 8

US-09-198-452A-1030
Sequence 1030, Application US/09198452A
Patent No. 6559294

GENERAL INFORMATION:
APPLICANT: Grifals, R.

TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection

TITLE OF INVENTION: and treatment of infection

FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/09/198, 452A

CURRENT FILING DATE: 1998-11-24

NUMBER OF SEQ ID NOS: 6849

SEQ ID NO 1030

LENGTH: 945

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

US-09-198-452A-1030

Query Match

Best Local Similarity 27.5%; Score 536; DB 4; Length 945;

Matches 116; Conservative 69; Mismatches 115; Indels 30; Gaps 6;

QY 32 LPKHIAQAMEELYQYKNAVL-QSAAPHAEIVLSNMRVAFDRMFLDVKEPEFSPVHEA 90
DB 22 LPPLQKSFVFNQITDIAATKAAADQAEVLQWQVIE-----DLKNPFIFFPYHK 77
QY 91 ILPEFNYMFGQYIRPLNPNRESYVGNVSFVQMEQLKQGDVVLISNHOEADPAVI 150
DB 78 IRAPIDLFRISIDFSLIDDKNSRLNLRLKLEIEYIARGDNVLLANHQTECDPQLM 137
QY 151 ALMLETNPHISNIIVAGDRVITDPLCKPFSMGRNLLCVYSKGMNDVPELAEMKKS 210

DB 138 YYALGKTHPELMENMIFVAGDRVITDPLARPFMSGCDLLCIYSKRHIATPEEREKILH 197
QY 211 NTRSLKEMALLRLRGSKIIWIAPSGGRDRPDPTNQWFPAPDTSLDNMRRLVDHAGLV 270
DB 198 NOKSMQILKTLNNEGKFIYVAPAGGRDRKN-AEGRLYPSEFSPESIEVFRLLAKASQT 256
QY 271 GHVYPLAILCHDIMPPLQVEKEIGKRLISHFGTGISVAPE-INFOEVTASCGSPPEAK 329
DB 257 THYPFALKTIDILPPPKKIENAIHQRAIFF-----APVFNFQ----- 296
QY 330 AAYSQALYDSVCEQYKVLHSAVHGKGLKLEA 359
DB 297 --AELFFDALCSKEELIHCDKHAQRTLRA 323

RESULT 9

US-09-134-001C-5549

Sequence 5549, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134, 001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064, 964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055, 779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 5549

LENGTH: 336

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5549

Query Match

Best Local Similarity 5.3%; Score 103.5; DB 4; Length 336;

Matches 49; Conservative 35; Mismatches 78; Indels 33; Gaps 9;

QY 49 NAVLOSA-APHAB-DIVLSNMRVAFDRMFLDVKEPEFSPVHEAI--LEPFNYMFGQNY 104
DB 137 NAVMEDGEIVHGETNIPKTHKI--DRVFL---ESDVEPMNEATEALEQADLIVLPG- 190
QY 105 IRPLNFRSYVGNVSFVQMEEQKQGDVVLISNHOEADPAVIALMLETNPHISEN 164
DB 191 -----SLYTSVISNLCVKGISEALLRTSAPKLYVSNVMTQPG-----ETDNVDVKEH 237
QY 165 IIVV---AGDRVITDPLCKPFSMGRNLLCVYSKGMNDVPELAEMKKSNTSLKEMALL 221
DB 238 IDALTRQVGEPIFDVFCSSYSKDVLRQYBEKSKPVAVHKEQLKDSGIRVLTASNIV 297
QY 222 -----LRGGSKII 229
DB 298 EISNEHYVHRNTKVL 312

RESULT 10

US-09-091-219-24

Sequence 24, Application US/09091219

Patent No. 6171592

GENERAL INFORMATION:

APPLICANT: STUDDERT, Michael J.

APPLICANT: CRABE, Brendan S.

APPLICANT: FENG, Li

TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS

FILE REFERENCE: 040268/0151

CURRENT APPLICATION NUMBER: US/09/091, 219

CURRENT FILING DATE: 1998-10-05

EARLIER APPLICATION NUMBER: PCT/AU96/00815

EARLIER FILING DATE: 1996-12-18

EARLIER APPLICATION NUMBER: AU PN7201

; EARLIER FILING DATE: 1995-12-18
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 2318
 ; TYPE: PRT
 ; ORGANISM: Foot-and-mouth disease virus
 US-09-091-219-24

Query Match 5.2%; Score 101.5; DB 3; Length 2318;
 Best Local Similarity 21.4%; Pred. No. 0.74; 137; Indels 103; Gaps 19;
 Matches 79; Conservative 50; Mismatches 137; Indels 103; Gaps 19;

QY 12 DARSEQDLGSGIORELE-----AGTLPKHI-----AQAMEELYQNYKN----- 49
 Db 857 ECRYNRAVNPLRGDLQVLAQKAVARTLPTSFNYGAIKATRTVELLYMKAEATYCPRPILL 916
 QY 50 AVLOSAPHAEDIVLS-NMRVAFDRMPL--DVKE---PFEFSPYH---EAILPEFNYYM- 99
 Db 917 AIHPTEARHKQKIVAPVKQTINFDLLKLAGDVESNPGFFFDVRSNFSKLIVETINQMOQE 976
 QY 100 -----FGONYIRPLVNFRESYVGVNVSFVGMEEQLKQGDVKVLISNHOSEA-----DP 147
 Db 977 DMSTKHGPDENLVSAFEELAIQVKAIRTGIDLEAKPWYKLIKLSRLSCMAVAARSKDP 1036
 QY 148 AVIALMETNPHISENIYVAGDRVITDPL-----CKPFSMGRNLLCV-----YSKXH 196
 Db 1037 VLVAIMLADTGLEILDSTFV---KKISDSLSLHFHPAPVPSFGAPVLLAGLVKVASF 1093
 QY 197 MNDVPE-LAEMKKRSNTRSLKEMALLRGSKII-----WIAP----- 233
 Db 1094 FRSTPEDLERAKQKARDINDIFAILKNGEWLVKILAIRDWIKAWIASSEKFFVTMTDL 1153
 QY 234 -----SGGRDRPDPI-----TNQWPPAPDATSLDNMRRLVDHAGLVGHYIPLAILCHDIM 284
 Db 1154 VPGILEKQDRLNDPSKYKEKEW-----LDNARQACLKSGNV-HI---ANLCKVVA 1200
 QY 285 PPPLOVEKE 293
 Db 1201 PAPSKSRPE 1209

RESULT 11

US-09-660-541-24
 ; Sequence 24, Application US/09660541
 ; Patent No. 6531136

; GENERAL INFORMATION:
 ; APPLICANT: STUDDERT, Michael J.
 ; APPLICANT: CRABB, Brendan S.
 ; APPLICANT: FENG, Li
 ; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
 ; FILE REFERENCE: 040268/0151
 ; CURRENT APPLICATION NUMBER: US/09/660,541
 ; CURRENT FILING DATE: 2000-09-12
 ; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/091,219
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 2318
 ; TYPE: PRT
 ; ORGANISM: Foot-and-mouth disease virus
 US-09-660-541-24

Query Match 5.2%; Score 101.5; DB 4; Length 2318;
 Best Local Similarity 21.4%; Pred. No. 0.74; 137; Indels 103; Gaps 19;
 Matches 79; Conservative 50; Mismatches 137; Indels 103; Gaps 19;

QY 12 DARSEQDLGSGIORELE-----AGTLPKHI-----AQAMEELYQNYKN----- 49
 Db 857 ECRYNRAVNPLRGDLQVLAQKAVARTLPTSFNYGAIKATRTVELLYMKAEATYCPRPILL 916

QY 50 AVLOSAPHAEDIVLS-NMRVAFDRMPL--DVKE---PFEFSPYH---EAILPEFNYYM- 99
 Db 917 AIHPTEARHKQKIVAPVKQTINFDLLKLAGDVESNPGFFFDVRSNFSKLIVETINQMOQE 976
 QY 100 -----FGONYIRPLVNFRESYVGVNVSFVGMEEQLKQGDVKVLISNHOSEA-----DP 147
 Db 977 DMSTKHGPDENLVSAFEELAIQVKAIRTGIDLEAKPWYKLIKLSRLSCMAVAARSKDP 1036
 QY 148 AVIALMETNPHISENIYVAGDRVITDPL-----CKPFSMGRNLLCV-----YSKXH 196
 Db 1037 VLVAIMLADTGLEILDSTFV---KKISDSLSLHFHPAPVPSFGAPVLLAGLVKVASF 1093
 QY 197 MNDVPE-LAEMKKRSNTRSLKEMALLRGSKII-----WIAP----- 233
 Db 1094 FRSTPEDLERAKQKARDINDIFAILKNGEWLVKILAIRDWIKAWIASSEKFFVTMTDL 1153
 QY 234 -----SGGRDRPDPI-----TNQWPPAPDATSLDNMRRLVDHAGLVGHYIPLAILCHDIM 284
 Db 1154 VPGILEKQDRLNDPSKYKEKEW-----LDNARQACLKSGNV-HI---ANLCKVVA 1200
 QY 285 PPPLOVEKE 293
 Db 1201 PAPSKSRPE 1209

RESULT 12

US-09-134-000C-4401
 ; Sequence 4401, Application US/09134000C
 ; Patent No. 6617156

; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR FILING DATE: 1997-08-15
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4401
 ; LENGTH: 426
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-4401

Query Match 5.0%; Score 97; DB 4; Length 426;
 Best Local Similarity 19.4%; Pred. No. 0.13;
 Matches 76; Conservative 57; Mismatches 123; Indels 136; Gaps 17;

QY 26 ELEAGTLPKHIAQAMEELYQNYKNVLSQAAPHAEDIVLSNMRV---AFDR-MFLDV-- 78
 Db 96 EKAGSLAPVAERSELVGPVTFVETRGKELEDAV--NNMKVGDVLVFENTREDVDG 154
 QY 79 -KE---PFEFSPYHEAILEPFNYMFGONYIRPLVNFRESYVGVNVSFVGMEEQL----- 129
 Db 155 KKEGNDALGKYWASLGDFVNDAGFAGTAHRAHASNVGIASGTGIPVAGFLMEKIKFG 214
 QY 130 -----KQGDVKVLISNHOSEADPAVIA-----LMLE 155
 Db 215 EAVENPKRPFVAILGGAKVSDKIAVIENTLIEKADKILIGGMAVTFMKAQGYSVGLSLE 274
 QY 156 TTNPHISENIYVAGDRVITDPLCKPFSMGRNLLCVSKGHNDVP-----ELAEMKK 208
 Db 275 EKVDLAKSLMEKAGDKVLV-----VDTVSKFSDNAPFHTVPTSTEIPDDEE 323
 QY 209 RSN--TRSLKEMALLRGSKIIWIAPSGGRDRPDPIITNQWPPAPDATSLDNMRRLVDH 266
 Db 324 GLDIGEKTIELFANELQAKTVVWNGPMG-----VFMSNFAK----- 361
 QY 267 AGLVGHYIPLAILCHDIMPPPLQVEKEIGERKLISFHGTGISVAPEINQFVTCGSPPE 326

```
Db 362 -GTIG-----VCEAI-----ANLEDAITIIIGGD 384
QY 327 EAKAAYSQALYDSYCEQYKVLHSAVHGKGLG 358
Db 385 SAAAI-QLGVN-----KFSHISTGGASIE 410

RESULT 13
US-09-328-352-7721
; Sequence 7721, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7721
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7721

Query Match 5.0%; Score 96.5; DB 4; Length 575;
Best Local Similarity 20.6%; Pred. No. 0.25;
Matches 68; Conservative 52; Mismatches 109; Indels 101; Gaps 14;

QY 112 RESYVGNVS-----FGVWERQLKQGDVKVLISNHOETADPAVIALMLETT- 157
Db 202 QESINGNAVVKFAGSESQERFYKSEENLRKGLKWIVQNLNFPVQVWACAMALIV 261
QY 158 -----NPHISENIIVAGDRVITDPLCKPFS-----MGRNLLCVYKHKHMDVPE 202
Db 262 WLALRPQILGNTTAGFVAVITAAGLLSKPV-KNLTDVNEKIQGLAAHSAHFELLDLPE 320
QY 203 ---LAEMKKR-----SNTSLKEMALLRGSKIIWIAPSGGRDRPDPT 244
Db 321 EONGSELKQLQGAIRFDHVLVNYADGTQAIKDFSLDIRPGETVALVGRSGA----- 372
QY 245 NQWFPAPPDATSLDNMRLVDHAGLV-----GHYPLAILCHDIMPPLQVKEIGEK 297
Db 373 -----GKTSLVNM-----LVRFQVSSGGIYLDLPDIRLELSLRTQIAWVNO 416
QY 298 RLISHFGTGISVAPENFOVTASCSPEE-----AKAAYSQALYDSYCEQYKVLHSAVHG 353
Db 417 QVVLFNRT---VRENIAYGQLHNA--SDEDVIAAKAAVAHDFIMNLPNGY-----DTVLG 467
QY 354 GKGLEAS-----TPSVLSLOPLD 374
Db 469 AGLNLSSGQORRIATARAILKNAPILILD 497

RESULT 14
US-09-312-283C-396
; Sequence 396, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: and Methods for Their Use
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 396
```

```
; LENGTH: 1529
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-396
```

```
Query Match 4.9%; Score 95; DB 4; Length 1529;
Best Local Similarity 24.0%; Pred. No. 1.9;
Matches 99; Conservative 54; Mismatches 143; Indels 116; Gaps 26;
```

```
QY 32 LPKHIAQAMEEL-YQNYKNAVLSQAA-----PHAEDIVLSNMVRVAFDRMFLDYKE-PFE- 83
Db 527 IPDHIPQYTAELRLNNEEPTVLEATGIFKKLPQLRKINLSNNKIT-----DIEEGAFEG 580
QY 84 FEPYHEAI-----LEFPNYMF-QQNYIRPLV--NFRSYVGNVSVFGVMEELQKQGDVK 135
Db 581 ASGVNEILTSNRLENVQHKMFGLSESLKTLMLRSNRISCVGNDSTGLGSVRL-----L 635
QY 136 VLISNHOETADPAVIALM--LET---TNP-----HISENIIVAGDRVIT-DPLC-KPF 182
Db 636 SLVDNQITTVAPGAGTGLHSLSTLNLANFPNCNCHLAWLGELRRKRIVTGNPRCKPY 695
QY 183 SMG-----RNLLCVYKHKHMDVPELAEMKKRSNTRSLKEMALL---LRGGSKIIM 230
Db 696 FLKEIPIQDVAIQDFTC-----DDGNDNSCSPL-----SRCPSECTCLDTVVVRCNKG 746
QY 231 IAPSGGRDRPDPT-----NOMFPAPFDAT-----SLDNMRL 263
Db 747 VLPKG---IPRVTEIYLDGNQFTLVPKELSNYKHLTLIDLNNRISTLSNOSFSNMTOL 803
QY 264 VDHAGLVGHYPLAILCHD---IMPPPLQVKEIGEKRLISPHGTGISVAPENFOEVT 320
Db 804 L-----TLISYNRLCIPP--RTFGLKSLRLSLHGNDISVVPFGAFCDLSA 850
QY 321 SCGSPEEAKAY-----SQALYDSYCEQYKVLHSAVHGKG-----LEASTPS 363
Db 851 LSHLAGANPLYCDCNMQLSDWVKSEYKEPGIARCAPGEMADKILLTTTPS 902
```

```
RESULT 15
US-09-252-991A-31268
; Sequence 31268, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31268
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31268

Query Match 4.8%; Score 93.5; DB 4; Length 453;
Best Local Similarity 19.5%; Pred. No. 0.35;
Matches 70; Conservative 59; Mismatches 131; Indels 99; Gaps 18;
```

```
QY 5 GHSRTFFDARSEQDILLSGIQRELEAGTLPKHIAQAMEELYQNYKNAVLSQAA----- 57
Db 127 GKSVRILEASDADELFSVILEYAKRLLENR--KKLQBIWEKTPMNVSEASEPVGALKKEE 184
QY 58 ----HAEDIVLSNMVRVAFDRMFLDYKEFPFESPVHEA-----LLE 93
Db 185 PVQSASQDVANGKLVADVAIAIEHAQONAASEYGEARRYLOVDFDNGSGSTDAVMTYLE 244
QY 94 PFNYMFQNYIRPLVNFRESYVG--NVSVFQVMEELQKQGDVK---VLISNHQTEAD-- 146
```

Search completed: March 16, 2004, 02:08:14
Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2004, 00:53:55 ; Search time 26 Seconds
(without alignment)

749.009 Million cell updates/sec

Title: US-09-926-805-7

Perfect score: 1949

Sequence: 1 GSMHGHSTFIDARSEQDLL.....KGLEASTPSVLSQPLQFLD 374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1913.5	98.2	463	1	PLSB CARTI
2	1389	71.3	459	1	PLSB ARATH
3	1379	70.8	457	1	PLSB PEA
4	1377	70.7	470	1	PLSB CUOSA
5	1350	69.3	396	1	PLSB CUCMO
6	1332	68.3	461	1	PLSB PHAVU
7	1321	67.8	472	1	PLSB SPIOL
8	103.5	5.3	882	1	PPDK RICCN
9	101.5	5.2	539	1	CH60 BACTR
10	101.5	5.2	2332	1	POLG FMDVO
11	98	5.0	2332	1	POLG FMDVA
12	97.5	5.0	545	1	CH60 OCEIH
13	96	4.9	432	1	SYH PYRFU
14	95.5	4.9	880	1	POLG RICPR
15	94	4.8	1061	1	ECB4 ARATH
16	93.5	4.8	539	1	CH60 BACST
17	93.5	4.8	604	1	SP20 YEAST
18	92.5	4.7	400	1	PGK LACPL
19	92.5	4.7	542	1	CH60 LISIN
20	92.5	4.7	2748	1	NUM1 YEAST
21	92	4.7	1529	1	SLR2 HUMAN
22	92	4.7	1571	1	ATG5 YEAST
23	91.5	4.7	448	1	TX19 HUMAN
24	91.5	4.7	543	1	CH60 CLOAB
25	91.5	4.7	684	1	CPSC HUMAN
26	91.5	4.7	810	1	SC12 ARATH
27	91.5	4.7	950	1	ATC1 YEAST
28	91	4.7	219	1	PIMT PYRAB
29	91	4.7	543	1	CH60 BACSU
30	91	4.7	1061	1	ECAL ARATH
31	91	4.7	2333	1	POLG FMDV1
32	91	4.7	2336	1	POLG FMDVZ
33	90.5	4.6	394	1	ALG2_PSEFL

34 90.5 4.6 542 1 CH60 LISMO
35 90.5 4.6 1376 1 RPOD ARATH
36 90 4.6 544 1 CH60 BACHD
37 89.5 4.6 684 1 CPSC BOVIN
38 89 4.6 783 1 NQO3 THERM
39 89 4.6 1690 1 RPOC THERM
40 89 4.6 2128 1 SPCB MOUSE
41 89 4.6 2183 1 DO10 HUMAN
42 88.5 4.5 332 1 LIPB KUULA
43 88.5 4.5 427 1 MURA BORBU
44 88.5 4.5 431 1 SYH PYRAB
45 88 4.5 1384 1 RPOD SINAI

ALIGNMENTS

RESULT 1
PLSB CARTI STANDARD; PRT; 463 AA.
AC Q42713;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycero1-3-phosphate acyltransferase, chloroplast precursor
DE (EC 2.3.1.15) (GPAT)
OS Carthamus tinctorius (Safflower)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Carduoideae; Cardueae; Carthamus.
OX NCBI_TaxID=4222;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RX MEDLINE=95148760; PubMed=7846182;
RA Bhella R.S., Mackenzie S.L.;
RT "Nucleotide sequence of a cDNA from Carthamus tinctorius encoding a
glycero1-3-phosphate acyl transferase.";
RL Plant Physiol. 106:1713-1714(1994).
CC -!- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
of glycerol-3-phosphate. The enzyme from chilling-resistant plants
discriminate against non-fluid palmitic acid and select oleic acid
whereas the enzyme from sensitive plants accepts both fatty acids.
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
acyl-sn-glycerol 3-phosphate.
CC -!- PATHWAY: First step in de novo phospholipid biosynthesis. It may
also function in the regulation of membrane biogenesis.
CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -!- SIMILARITY: Belongs to the plant GPAT family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L33841; AAA74319.1;
CC InterPro: IPR002123; Acyltransferase.
CC Pfam: PF01553; Acyltransferase; 1.
CC SMART: SM00563; PlsC; 1.
CC KW Phospholipid biosynthesis; Transferase; Acyltransferase;
FT Transit peptide; Chloroplast.
FT CHAIN 1 91 CHLOROPLAST (POTENTIAL).
FT TRANSIT 92 463 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
SQ SEQUENCE 463 AA; 50834 MW; 5792E933068A534D CRC64;

Query Match 98.2%; Score 1913.5; DB 1; Length 463;
Best Local Similarity 98.7%; Pred. No. 1.5e-150;
Matches 370; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
1 GSMHGHSTFIDARSEQDLLSGIQRELEAGTLPKHIAQAMEELYQNYKNAVLSAAPHAE 60

Db 89 GSDHGHSTFIDARSEQDLSGIORELEAGTLPKHIAQMBELYQNVKAVLQSAAPHAE 148
 QY 61 DIVLSNMRVAFDRMFLDKPEFSPHYEALLESFNYMFGQNYRPLVNPRESYGVNS 120
 Db 149 DIVLSNMRVAFDRMFLDKPEFSPHYEALLESFNYMFGQNYRPLVNPRESYGVNS 208
 QY 121 VFGWEEQLKQDKVLSNHNQTEADPAVIALMETNPHISENIIVAGDRVITDPLCK 180
 Db 209 VFGWEEQLKQDKVLSNHNQTEADPAVIALMETNPHISENIIVAGDRVITDPLCK 268
 QY 181 PFSMGRNLLCVYKXKNDVDELAKMKRSNTRSLK-EMALLRGSKIIWIAPSGGRDR 239
 Db 269 PFSMGRNLLCVYKXKNDVDELAKMKRSNTRSLKGRMALLRGSKIIWIAPSGGRDR 328
 QY 240 PDPITNWFPAFPDTSLDNNRRLVDHAGLVGHYPLAILCHDIMPPLQVEKEIGKRL 299
 Db 329 PDPITNWFPAFPDTSLDNNRRLVDHAGLVGHYPLAILCHDIMPPLQVEKEIGKSW 388
 QY 300 ISFHTGTSISVAPEINQFVTTASCSPPEAKAAYSQALYDSYCEQYKVLHSAVHGKGLEA 359
 Db 389 ISFHTGTSISVAPEINQFVTTASCSPPEAKAAYSQALYDSYCEQYKVLHSAVHGKGLEA 448
 QY 360 STPSVLSQPLQFLD 374
 Db 449 STPSVLSQPLQFLD 463

RESULT 2

PLSB ARATH STANDARD; PRT; 459 AA.
 AC Q43307; GQFV85;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
 DE (EC 2.3.1.15) (GPAT).
 GN AT51 OR AT1G32200 OR F3C3.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93144702; PubMed=7678766;
 RA Nishida I., Tasaka Y., Shiraiishi H., Murata N.;
 RT "The gene and the RNA for the precursor to the plastid-located
 RT glycerol-3-phosphate acyltransferase of Arabidopsis thaliana.";
 RL Plant Mol. Biol. 21:267-277(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.P., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu X.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Militscher J., Miranda M., Nguyen M., Nietman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).

103
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J., Cheuk R.F.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Yu G.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowsher L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallenberg E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:842-846(2003).
 CC -1- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
 CC of glycerol-3-phosphate. The enzyme from chilling-resistant plants
 CC discriminate against non-fluid palmitic acid and select oleic acid
 CC whereas the enzyme from sensitive plants accepts both fatty acids.
 CC This is an oleate-selective acyltransferase.
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 CC acyl-sn-glycerol 3-phosphate.
 CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 CC also function in the regulation of membrane biogenesis.
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -1- SIMILARITY: Belongs to the plant GPAT family.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; D00673; BAA00576.1; -;
 DR ENBL; D00672; BAA00575.1; -;
 DR ENBL; AC084165; AAG23437.1; -;
 DR ENBL; AY093169; AAM13168.1; -;
 DR ENBL; BT008758; AAP49520.1; -;
 DR PIR; E86446; E86446.
 DR PIR; S31089; S31089.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 DR SMART; SM00563; Pfam; 1.
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;
 KW Transit peptide; Chloroplast.
 FT TRANSIT 1 90 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 91 459 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
 FT CONFLICT 93 93 S -> N (IN REF. 1).
 FT CONFLICT 287 287 D -> V (IN REF. 1).
 SQ SEQUENCE 459 AA; 50421 MW; 684CF97EASB82A7E CRC64;
 Query Match 71.3%; Score 1389; DB 1; Length 459;
 Best Local Similarity 69.5%; Pred. No. 3.5e-107;
 Matches 253; Conservative 57; Mismatches 54; Indels 0; Gaps 0;
 QY 6 H8RTFDARSEQDLSGIORELEAGTLPKHIAQMBELYQNVKAVLQSAAPHAE 65
 Db 94 H8RTFDARSEQDLSGIGKEAEAGRLPANVAGHEELYWNVKNVLSGASRAETVVS 153
 QY 66 NMRVAFDRMFLDKPEFSPHYEALLESFNYMFGQNYRPLVNPRESYGVNS 125
 Db 154 NMSVAFDRMFLDKPEFSPHYEALLESFNYMFGQNYRPLVNPRESYGVNS 213
 QY 126 EQQLKQDKVLSNHNQTEADPAVIALMETNPHISENIIVAGDRVITDPLCKPFG 185
 Db 214 EDKIQGHNVLLISNHSQSEADPAVLSLLEAQSPFIGNIKCVAGDRVITDPLCKPFG 273

186 RNLICVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 245
 274 RNLICVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 333
 246 QWFPAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 305
 334 EWFAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 393
 306 GISVAPEINFEQVTAAGSPPEAKAAYSOALYDSVCEQYKVLHSAVHGKGLASTPSVS 365
 394 GLSIAPEINFEQVTAAGSPPEAKAAYSOALYDSVCEQYKVLHSAVHGKGLASTPSVS 453
 366 LSQP 369
 454 LSQP 457

Query Match 70.8%; Score 1379; DB 1; Length 457;
 Best Local Similarity 71.1%; Pred. No. 2.3e-106; Indels 0; Gaps 0;
 Matches 258; Conservative 53; Mismatches 52

QY 7 SRTFDARSEQDLSGIGRELEAGTLPHKIAQAAMEELYQYKNAVLQSAAPHAEDIVLSN 66
 Db 93 SRTFLNAQNEQDVLGIGKEVEAGTLPASIAAGMEEVLYNKSAVIGKSDPKANEIVLSN 152
 QY 67 MRVAFDRMFLDVKPEPEFSPYHEALILEPENYMFQNYIRPLVNFRESYVGVNVE 126
 Db 153 MYALLDRIFLDVKPEPVFAHHKAKREPPDYMFQNYIRPLVDFETSVGNMPLFIQME 212
 QY 127 EQLKQDKVLLISNHQTEADPAVIALMLTTPPHISENIYVAGDRVITDPLCKPFSMGR 186
 Db 213 EQLKQGHNILLNSHQSEADPAIALLLEMLPHIAENLIYVAGDRVITVPLCKPFSIGR 272
 QY 187 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 246
 Db 273 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 332
 QY 247 WFPAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 306
 Db 333 WAPAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 392
 QY 307 ISVAPEINFEQVTAAGSPPEAKAAYSOALYDSVCEQYKVLHSAVHGKGLASTPSVS 366
 Db 393 ISTAPEISFNTAACENPERAKADYTRALYDSVTEQYDVLKSAHKGKGLASTPSVS 452
 QY 367 SQP 369
 Db 453 SQP 455

RESULT 4
 PLSB CUCA
 ID PLSB CUCA STANDARD; PRT; 470 AA.
 AC Q39639;
 DT 15-DEC-1998 (Rel. 37; Created)
 DT 15-DEC-1998 (Rel. 37; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
 DE (EC 2.3.1.15) (GPAT).
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID:3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson T.C., Schneider J.C., Somerville C.;
 RT "Nucleotide sequence of acyl-acyl carrier protein: glycerol-3-
 phosphate acyltransferase from cucumber.";
 RL Plant Physiol. 99:771-772(1992).
 CC -!- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
 of glycerol-3-phosphate. The enzyme from chilling-resistant plants
 discriminate against non-fluid palmitic acid and select oleic acid
 whereas the enzyme from sensitive plants accepts both fatty acids.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 acyl-sn-glycerol 3-phosphate.
 CC -!- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 also function in the regulation of membrane biogenesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: Belongs to the plant GPAT family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC
 DR EMBL; X59041; CAA41769.1; -
 DR F01; S18239; S18239.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.

186 RNLICVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 245
 274 RNLICVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 333
 246 QWFPAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 305
 334 EWFAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 393
 306 GISVAPEINFEQVTAAGSPPEAKAAYSOALYDSVCEQYKVLHSAVHGKGLASTPSVS 365
 394 GLSIAPEINFEQVTAAGSPPEAKAAYSOALYDSVCEQYKVLHSAVHGKGLASTPSVS 453
 366 LSQP 369
 454 LSQP 457

Query Match 70.8%; Score 1379; DB 1; Length 457;
 Best Local Similarity 71.1%; Pred. No. 2.3e-106; Indels 0; Gaps 0;
 Matches 258; Conservative 53; Mismatches 52

QY 7 SRTFDARSEQDLSGIGRELEAGTLPHKIAQAAMEELYQYKNAVLQSAAPHAEDIVLSN 66
 Db 93 SRTFLNAQNEQDVLGIGKEVEAGTLPASIAAGMEEVLYNKSAVIGKSDPKANEIVLSN 152
 QY 67 MRVAFDRMFLDVKPEPEFSPYHEALILEPENYMFQNYIRPLVNFRESYVGVNVE 126
 Db 153 MYALLDRIFLDVKPEPVFAHHKAKREPPDYMFQNYIRPLVDFETSVGNMPLFIQME 212
 QY 127 EQLKQDKVLLISNHQTEADPAVIALMLTTPPHISENIYVAGDRVITDPLCKPFSMGR 186
 Db 213 EQLKQGHNILLNSHQSEADPAIALLLEMLPHIAENLIYVAGDRVITVPLCKPFSIGR 272
 QY 187 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 246
 Db 273 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEWALLLRGSKTIWIAPSGRRDRPDITN 332
 QY 247 WFPAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 306
 Db 333 WAPAPFADTSLDMRRLVDHAGLVGHYPLAILCHDIMPPLQVKEKEKELISFHGT 392
 QY 307 ISVAPEINFEQVTAAGSPPEAKAAYSOALYDSVCEQYKVLHSAVHGKGLASTPSVS 366
 Db 393 ISTAPEISFNTAACENPERAKADYTRALYDSVTEQYDVLKSAHKGKGLASTPSVS 452
 QY 367 SQP 369
 Db 453 SQP 455

RESULT 4
 PLSB CUCA
 ID PLSB CUCA STANDARD; PRT; 470 AA.
 AC Q39639;
 DT 15-DEC-1998 (Rel. 37; Created)
 DT 15-DEC-1998 (Rel. 37; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
 DE (EC 2.3.1.15) (GPAT).
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID:3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson T.C., Schneider J.C., Somerville C.;
 RT "Nucleotide sequence of acyl-acyl carrier protein: glycerol-3-
 phosphate acyltransferase from cucumber.";
 RL Plant Physiol. 99:771-772(1992).
 CC -!- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
 of glycerol-3-phosphate. The enzyme from chilling-resistant plants
 discriminate against non-fluid palmitic acid and select oleic acid
 whereas the enzyme from sensitive plants accepts both fatty acids.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 acyl-sn-glycerol 3-phosphate.
 CC -!- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 also function in the regulation of membrane biogenesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: Belongs to the plant GPAT family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC
 DR EMBL; X59041; CAA41769.1; -
 DR F01; S18239; S18239.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.


```

RESULT 6
PLSB PHAVU
ID PLSB PHAVU STANDARD; PRT; 461 AA.
AC Q43822;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
DE (EC 2.3.1.15) (GPAT).
GN PLSB.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OC NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Annabel; TISSUE=Leaf;
RX MEDLINE=95232196; PubMed=7716242;
RA Fritz M., Heinz E., Wolter F.P.;
RT "Cloning and sequencing of a full-length cDNA coding for
RT sn-glycerol-3-phosphate acyltransferase from Phaseolus vulgaris.";
RL Plant Physiol. 107:1039-1040(1995).
CC -1- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
CC of glycerol-3-phosphate. The enzyme from chilling-resistant plants
CC discriminates against non-fluid palmitic acid and select oleic acid
CC whereas the enzyme from sensitive plants accepts both fatty acids.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may
CC also function in the regulation of membrane biogenesis.
CC -1- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -1- SIMILARITY: Belongs to the plant GPAT family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X79722; CAA56159.1; --
CC PIR; T11819; T11819.
CC InterPro; IPR002123; Acyltransferase.
CC Pfam; PF01553; Acyltransferase; 1.
CC SMART; SM00563; PlsC; 1.
CC KX Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Transit peptide; Chloroplast.
CC FT TRANSIT 1 96
CC CHAIN 97 461 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
CC SEQUENCE 461 AA; 50697 MW; EAC4FC937908B38A CRC64;
Query Match
Best Local Similarity 67.2%; Pred. No. 1.8e-102;
Matches 244; Conservative 61; Mismatches 58; Indels 0; Gaps 0;
QY 7 SRTFDARSEQDLGSIQRELEAGLTPKHIAQAMEELYCNKYNALVQSAAPHAEIVLSN 66
Db 97 SRTFLNAQSEQDVFAKKEVEAGSLPANVAGMEEVNKNYKAVIQSGDPRANEIVLSN 156
QY 67 MRVAFDRFLVDKPEFSPVHEALEPFFNYMFQNYIRPLNFRSIVGNVSVFGVME 126
Db 157 MIALDRVFLVDTPDFVPHKAKREPFDYVVFQNYIRPLNVPKNAVYGNMPLFIEME 216
QY 127 EQLKGDKGVLLSNHQTADPAVIALMETTNPHSENIIVAGDRVITDPLCKPFGMR 186
Db 217 EKLKGHNHLLMSNHQTADPAISLLETLPLPYIAENLTIVAGDRVITDPLSKPFSIGR 276
QY 187 NLLCVYSKHHNDVPELAEMKRSNTRSLKEMALLRGSKIWIAPSGGRDRPDPTNQ 246
Db 277 NLCVYSKHHMLDDPALVEMKRTANIRALKEMALLRNGSLQVWVIAPSGGRRDPAQTR 336

```

us-09-926-805-7.rsp

Page 5

RESULT 7

PLSB SPIOL

ID PLSB SPIOL STANDARD; PRT; 472 AA.

AC Q43869;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glycerol-3-phosphate acyltransferase, chloroplast precursor

DE (EC 2.3.1.15) (GPAT).

GN GAT OR ACT1.

OS Spinacia oleracea (Spinach).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllales; Amaranthaceae; Spinacia.

OC NCBI_TaxID=3562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Birefore; TISSUE=Seedling;

RX Ishikazi-Nishizawa O., Azuma M., Ohtani T., Murata N., Toguri T.;

RT "Nucleotide sequence of cDNA from Spinacia oleracea encoding plastid

RT glycerol-3-phosphate acyltransferase.";

RL (In) Plant Gene Register PGR95-014.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Melody; TISSUE=Leaf;

RX Wolter F.P.;

RT "Conserved intron position in 3' untranslated region of a cDNA

RT encoding the plastidial sn-glycerol-3-phosphate acyltransferase of

RT spinach.";

RL (In) Plant Gene Register PGR96-118.

RN [3]

RP CHARACTERIZATION.

RX MEDLINE=83131613; PubMed=6825679;

RA Frentzen M., Heinz E., McKeon T.A., Stumpf P.K.;

RT "Specificities and selectivities of glycerol-3-phosphate

RT acyltransferase and monoacylglycerol-3-phosphate acyltransferase from

RT pea and spinach chloroplasts.";

RL Eur. J. Biochem. 129:629-636(1983).

CC -1- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position

CC of glycerol-3-phosphate. The enzyme from chilling-resistant plants

CC discriminates against non-fluid palmitic acid and select oleic acid

CC whereas the enzyme from sensitive plants accepts both fatty acids.

CC This is an oleate-selective acyltransferase.

CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-

CC acyl-sn-glycerol 3-phosphate.

CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may

CC also function in the regulation of membrane biogenesis.

CC -1- SUBCELLULAR LOCATION: Chloroplast stroma.

CC -1- SIMILARITY: Belongs to the plant GPAT family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; X77370; CAA54559.1; --

CC EMBL; Z49091; CAA88913.1; --

48

```

QY 237 -----RD-----PITNMFAPFAPDATSLDNMR-RLVDHAGLVGHY 274
Db 578 NKIPLVDMITAPDIERRKLAQVQLLPLOYEDFVALFRVMDKPVNRLDLP----- 629
QY 275 PLALCHDIMPPLQVKEIG-----EKLSIFHGIGISVAPEINFQEVTA SCGS 324
Db 630 PL-----HEFLPTTEEDKKNLSNLSNPLSMINQRLHAMH-----EVNPMLGHRGCRIGCS 681
QY 325 PEEAKAAYS---QALYDSVCEQYKVLHSAVHGKGLASSTPSVS 365
Db 682 PE---LYQWQIEAIFTAIFE---LHKXGHEICNLEIMPLIS 717

RESULT 9
ID CH60 BACTR STANDARD; PRT; 539 AA.
AC QRVV84;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN 60 kDa chaperonin (Protein Cpn60) (groEL protein).
OS Bacillus thermoglucosidasius (Geobacillus thermoglucosidasius).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1426;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=KP1006;
RX MEDLINE=21823396; PubMed=11834128;
RA Watanabe K., Fujiwara H., Inui K., Suzuki Y.;
RT "oligo-1,6-glucosidase from a thermophile, Bacillus
thermoglucosidasius KP1006, was efficiently produced by combinatorial
expression of GroEL in Escherichia coli.";
RL Biotechnol. Appl. Biochem. 35:35-43(2002).
CC -!- FUNCTION: Prevents misfolding and promotes the refolding and
proper assembly of unfolded polypeptides generated under stress
conditions (By similarity).
CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
7 subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; AB025944; BAB83940.1; -
DR HAMAP; MF 00600; -; 1.
DR InterPro; IPR001844; Chaprin Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR Pfam; PF00118; cpn60 TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding.
SQ SEQUENCE 539 AA; 57192 MW; 70GBE183F07C67B5 CRC64;

Query Match 5.2%; Score 101.5; DB 1; Length 539;
Best Local Similarity 19.3%; Pred. No. 1.2;
Matches 91; Conservative 63; Mismatches 131; Indels 187; Gaps 20;

QY 22 GQRELEAGTIPKHTAQAMELYQ-----NYKNAVLSQAAPHAED-----IVLSNMRVA 70
Db 114 GIKGIE-----KAVAAVEELKALSKPIQKESIAQVAALSADEEVGLIAEMERVG 168
QY 71 FDRMF-----LDVKEPFEF-----SPY-----HEALEPPNYMFGQNYRP 107

```

```

Db 169 NDGVITLESKGTTELDVVEGMOFGYASPYMITDEKAEAVLE-----NP 216
QY 108 LVNFRESYGVNV-SVFGVMEELKQGDVKVLI SNH-QTEADPAVIALMLTETNPHS--- 162
Db 217 YILITDKISNIQDILPILQVQVQCKPLLIADVEGEALATVYVVKLGRGTTAVAVKA 276
QY 163 -----ENIYVAGDRVITDPLKPF-----SMGRNLLCVYSKKH--- 196
Db 277 PGFGRRKAMLEIAITLGTGEVISEELGRELKLTITIASLGASKVVTVENTTIVEGAGD 336
QY 197 -----MNDV-----PELAEMKKR--- 209
Db 337 SERIKARINQIRAOLETTSEFDRGKLOERLAKLGGVAVIKVGAATETELKERKRIED 396
QY 210 ---SNTSLKEMALLRLGG-----SKIWIAPSGGRDRPDPTNWFAPFAPDATSLDNM 260
Db 397 ALNSTRAAVEEGIVAGGTTALNNVYNKVAIEAEGDEATCVKIVLRAIEP-----V 448
QY 261 RLVDHAGLVGHYPLAIIICHDIMPPLQVKEIGKRLISFHGTGISVAPEINFQEVTA 320
Db 449 RQIAQVAGLEGSVI-----VERLKSEK-----PGIGFNAATG 480
QY 321 SCGSPEEAKAAYSQALYDSVCEQYKVLHSAVHGKGLER---STPSVSLSQP 369
Db 481 EWNVMEIA-----GIVPTKVRSAQNAAASVAAMFLTTEAVVADKP 522

RESULT 10
POLG FMDVQ
ID POLG FMDVQ STANDARD; PRT; 2332 AA.
AC P03305;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Genome polypeptide [Contains: Nonstructural protein P20A; Coat
protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;
Core protein P12; Core protein P34; Core protein P14; Genome-linked
protein VP6; Protease (EC 3.4.22.-); RNA-directed RNA polymerase
(EC 2.7.7.48)].
DE Foot-and-mouth disease virus (strain O1) (Aphthovirus O) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=73482;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=O1K;
RX MEDLINE=84297249; PubMed=6089122;
RA Forss S., Strebel K., Beck E., Schaller H.;
RT "Nucleotide sequence and genome organization of foot-and-mouth
disease virus.";
RL Nucleic Acids Res. 12:6587-6601(1984).
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=O1BFS;
RX MEDLINE=83143292; PubMed=6298715;
RA Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;
RT "Comparison of the amino acid sequence of the major immunogen from
three serotypes of foot and mouth disease virus.";
RL Nucleic Acids Res. 10:8285-8295(1982).
RN [3]
RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=89143740; PubMed=2537470;
RA Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;
RT "The three-dimensional structure of foot-and-mouth disease virus at
2.9-A resolution.";
RL Nature 337:709-716(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN.

```


Db 401 TRAGVEGMEVGGGTAIVNIHQVSELTEGDEATGASIVLRALREPRQIVHNAGLEGS 460
 Qy 273 IYPLAILCHDIMPPLQVEKEIGKRLISFHTGTSVAPE--INFOEVTASCGSPPEAKA 330
 Db 461 II-----VERIKGEK-----VGIGYNAATDEWNNVEA----- 488
 Qy 331 AYSQALYSVCEQYKVLHSAVHGKGLEA---STPSVLSOP 369
 Db 489 -----GIVDPKTVRSALQNARASVAAMFLITTEAVVADKP 522

RESULT 13

SYN_PIRFU
 ID SYN_PIRFU STANDARD; PRT; 432 AA.
 AC Q8U431;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Histidyl-tRNA synthetase (SC 6.1.1.21) (Histidine--tRNA ligase)
 DE (HARS).
 GN HISS OR PF0264.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RX Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RA "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -! CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
 CC diphosphate + L-histidyl-tRNA(His).
 CC -! SUBCELLULAR LOCATION: Cytoplasmic.
 CC -! SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AE010150; AAL80398.1;
 CC HAMAP; MF 00127; -; 1.
 CC InterPro; IPR004154; HGTP anticodon.
 CC InterPro; IPR004516; Hiss-
 CC InterPro; IPR002314; tRNA-synt 2b.
 CC InterPro; IPR006195; tRNA_ligase II.
 CC Pfam; PF03129; HGTP anticodon; 1.
 CC Pfam; PF00587; tRNA-synt 2b; 1.
 CC TIGRFAMs; TIGR00442; hiss; 1.
 CC PROSITE; PS50862; AA tRNA_LIGASE II; 1.
 CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; Atp-binding;
 KW Complete proteome.
 SQ SEQUENCE 432 AA; 49183 MW; F6C8B0AD0BE150F CRC64;

Query Match 4.9%; Score 96; DB 1; Length 432;
 Best Local Similarity 21.4%; Pred. No. 2.5/
 Matches 61; Conservative 44; Mismatches 100; Indels 80; Gaps 14;
 Qy 66 NMVAFDRMELDKVEPFEPFSPYHEALEPFFNYMFGQYIRPLVNFRESYVGNVSGVM 125
 Db 77 DMTSSVRLVYNNM---FQTAP-----KPIKWY-----YIANMFRYEQSGRYRFPWA 122
 Qy 126 EEOLKQDKVLLSNHQTEDADPAVIALMLETTPHISENIYVA-----GDRVIT 175
 Db 123 GVELIGSDKY-----EADAEVIALFVES-----YLATGLRFTFVNIIGDRILL 164
 Qy 176 DPLCKPFSMGRN--LLCVYSKHMNDVPPELAEMKRSNTRSLKEMALLRGSKIIWIAP 233

Db 165 DEFAMLVGVDDIGIMRIIDKKO-----KLPQEDFINALKEFGLDKNGIEKVELEIN 216
 Qy 234 SGGSDRED---PITNFFPAPFDATSLDNMRRLVD-----HAGLV-CHYVPL 276
 Db 217 IKG--KPEVLPPLAEELFTSEVAKNEINRLVALIDLEAVEYKVDWIDJGIARGFDYTT 274
 Qy 277 ALLCHDIMPPLQVEKEIGKR---LISFHG-----TGISVAPE 312
 Db 275 SIVFEATAPNDLGIGSIGGGGRYDNLIEVFGKPTPATGFAIGIE 319

RESULT 14

PODK_RICPR
 ID PODK_RICPR STANDARD; PRT; 880 AA.
 AC Q9ZDS5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pyruvate,phosphate dikinase (EC 2.7.9.1) (Pyruvate,orthophosphate
 DE dikinase)
 DE PPK OR RP492.
 GN Rickettsia prowazekii.
 OS Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria";
 RL Nature 396:133-140(1998).
 CC -! FUNCTION: Catalyzes the reversible phosphorylation of pyruvate and
 CC phosphate (By similarity).
 CC -! CATALYTIC ACTIVITY: ATP + pyruvate + phosphate = AMP +
 CC phosphoenolpyruvate + diphosphate.
 CC -! COFACTOR: Magnesium ion (By similarity).
 CC -! SUBUNIT: Homodimer (By similarity).
 CC -! DOMAIN: The N-terminal domain contains the ATP/pi active site, the
 CC central domain the pyrophosphate/phosphate carrier His-460, and
 CC the C-terminal domain the pyruvate active site.
 CC -! MISCELLANEOUS: The reaction takes place in three steps, each
 CC mediated by a carrier histidine residue located on the surface of
 CC the central domain. The two first partial reactions are catalyzed
 CC at an active site located on the N-terminal domain, and the third
 CC partial reaction is catalyzed at an active site located on the C-
 CC terminal domain. For catalytic turnover, the central domain
 CC swivels from the concave surface of the N-terminal domain to that
 CC of the C-terminal domain.
 CC -! SIMILARITY: Belongs to the PEP-utilizing enzyme family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AJ235272; CAA14944.1;
 CC PIR; F71652; F71652.
 CC HSSP; P22983; 2DIN.
 CC InterPro; IPR008279; PEP mobile.
 CC InterPro; IPR000121; PEP utilizers.
 CC InterPro; IPR002192; PPK N term.
 CC Pfam; PF00391; PEP-utilizers; 1.
 CC Pfam; PF02896; PEP-utilizers; 1.
 CC Pfam; PF01326; PPK N term; 1.
 CC ProDom; PD000940; PEP utilizers; 1.
 CC PROSITE; PS00370; PEP_ENZYMES_PHOS_SITE; 1.

DR	PROSITE; PS00742; PEP_ENZYMES_2; 1.
KW	Transferase; Kinase; ATP-binding; Magnesium; Phosphorylation;
KW	Complete proteome.
FT	DOMAIN 1 348 N-TERMINAL.
FT	DOMAIN 349 405 LINKER 1.
FT	DOMAIN 406 503 CENTRAL.
FT	DOMAIN 504 538 LINKER 2.
FT	DOMAIN 539 880 C-TERMINAL.
FT	ACT_SITE 460 TELE-PHOSPHOHISTIDINE INTERMEDIATE
FT	(BY SIMILARITY).
FT	NP_BIND 97 97 ATP (POTENTIAL).
FT	NP_BIND 246 246 ATP (POTENTIAL).
FT	METAL 329 329 MAGNESIUM 1 (PROBABLE).
FT	METAL 331 331 MAGNESIUM 1 (PROBABLE).
FT	METAL 343 343 MAGNESIUM 1 (PROBABLE).
FT	METAL 750 750 MAGNESIUM 2 (PROBABLE).
FT	METAL 774 774 MAGNESIUM 2 (PROBABLE).
FT	MOD_RES 458 458 PHOSPHORYLATION (BY SIMILARITY).
SQ	SEQUENCE 860 AA; 98302 MW; C8CB877B7EAF3 CRC64;

Query Match	
Best Local Similarity 4.9%; Score 95.5; DB 1; Length 880;	
Matches 81; Conservative 77; Mismatches 131; Indels 193; Gaps 22;	

Qy	12 DARSEDLGSIQRELEAGTLPKHIAQAWEELY-----ONYKNAVLSQAAPHAED 61
Dd	298 DNSNQAMMPVEFKL-----SQAKKLSEHYLDMDQIEFTENNKLVIQLTRAKETA 351
Qy	62 IVLSNMVAFDRMFLDYKEPFEPSPYHEAIL---EPFNYYMFQG-NYIRPLVNFRSYV 116
Dd	352 IAINIAVQMVEKLISKE-----QALMRIDFESLNQLLTRIDYSKGLTSIAEGLP 403
Qy	117 GN-----VSFVGME-EOLKQGDQVLNISHQTEADPAVIALMETTNPHISENIIVYA 169
Dd	404 ASPGAATGIUVFSPYDAEKLSHHHKVILV-RHDTSP-----DINGMVSSGIITIR 454
Qy	170 GD-----RVITDPLCKFPFGRNLLCVYSKGH-----MND 199
Dd	455 GGMTSHAAVAVARGMKPCVCGTNNLSIDEQKQILIAGDIVIKQGDIIITDGSGKIFGE 514
Qy	200 VP-----ELAEPMKRSNTRSLEKEMALLRGSKLIW-----230
Dd	515 MPLIQPTFEBSKILLWADEISLKVRANAETVDNALSVIKFGAQIGLCRSEHMFFDK 574
Qy	231 -----IAPSGGR-----DRP-----DP'TNOMFWPA 250
Dd	575 NKIPLVREMIAPDIERRQCALKPLQTEDFKSLFRVMKNKPVTNLDDPPHLEFLP- 633
Qy	251 PFDATSLDNWRVLVDHAGLVGHVYPLAIL-----CHDIMPPPLQVEKEIGEKLISFHG 304
Dd	634 -----TTEEDKKNLANSNL-----PLSMIHQRHLAMEHVNP-----MLGHRG 671
Qy	305 TGISVA-PEINFOBVTTASCGSPPEAKAAYSQALYDSVCEQKVLSHAVHGKGLEGASTPS 363
Dd	672 CRLGICLPIFYQMCI-----EAIFTAIFE-----LHKKEHTIESNLELMIP 712
Qy	364 VS 365
Dd	713 IS 714

RESULT 15	
ECA4_ARATH STANDARD; PRT; 1061 AA.	
ID	ECA4_ARATH Q9XES1; Q9LOP2;
AC	Q9XES1; Q9LOP2;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Calcium-transferring AtPase 4, endoplasmic reticulum-type
DE	(SC 3.6.3.8).
GN	ECA4 OR AT1G07670 OR F24B9.24.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

FT	TRANSMEM	71	91	POTENTIAL.
FT	DOMAIN	92	115	LUMENAL (POTENTIAL).
FT	TRANSMEM	116	135	POTENTIAL.
FT	DOMAIN	136	278	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	279	298	POTENTIAL.
FT	DOMAIN	299	327	LUMENAL (POTENTIAL).
FT	TRANSMEM	328	345	POTENTIAL.
FT	DOMAIN	346	786	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	787	806	POTENTIAL.
FT	DOMAIN	807	816	LUMENAL (POTENTIAL).
FT	TRANSMEM	817	837	POTENTIAL.
FT	DOMAIN	838	857	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	858	880	POTENTIAL.
FT	DOMAIN	881	950	LUMENAL (POTENTIAL).
FT	TRANSMEM	951	970	POTENTIAL.
FT	DOMAIN	971	983	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	984	1002	POTENTIAL.
FT	DOMAIN	1003	1017	LUMENAL (POTENTIAL).
FT	TRANSMEM	1018	1038	POTENTIAL.
FT	DOMAIN	1039	1061	CYTOPLASMIC (POTENTIAL).
FT	MOD RES	383	383	PHOSPHORYLATION (BY SIMILARITY).
FT	METAL	731	731	MAGNESIUM (BY SIMILARITY).
FT	METAL	735	735	MAGNESIUM (BY SIMILARITY).
FT	METAL	336	336	CALCIUM 2 (VIA CARBONYL OXYGEN) (BY SIMILARITY).
FT	METAL	337	337	CALCIUM 2 (VIA CARBONYL OXYGEN) (BY SIMILARITY).
FT	METAL	339	339	CALCIUM 2 (VIA CARBONYL OXYGEN) (BY SIMILARITY).
FT	METAL	341	341	CALCIUM 2 (BY SIMILARITY).
FT	METAL	797	797	CALCIUM 1 (BY SIMILARITY).
FT	METAL	800	800	CALCIUM 2 (BY SIMILARITY).
FT	METAL	825	825	CALCIUM 1 (BY SIMILARITY).
FT	METAL	828	828	CALCIUM 2 (BY SIMILARITY).
FT	METAL	829	829	CALCIUM 1 (BY SIMILARITY).
FT	METAL	829	829	CALCIUM 2 (BY SIMILARITY).
FT	METAL	961	961	CALCIUM 1 (BY SIMILARITY).
FT	CONFLICT	314	314	P -> H (IN REF. 2).
FT	CONFLICT	430	430	T -> M (IN REF. 2).
FT	CONFLICT	439	439	M -> S (IN REF. 2).
FT	CONFLICT	659	659	T -> I (IN REF. 2).
FT	CONFLICT	859	859	T -> P (IN REF. 2).
FT	CONFLICT	1046	1046	G -> V (IN REF. 2).
SQ	SEQUENCE	1061 AA;	116180 MW;	55B6126ESD539822 CRC64;

Query Match
Best Local Similarity 4.8%; Score 94; DB 1; Length 1061;
Matches 54; Conservative 22.4%; Pred. No. 13;
Mismatch 45; Mismatches 106; Indels 36; Gaps 10;

QY	11	IDARSEQDILL---	SGIQRELEAGTLPKHIAQAMEELYQYKNVAVLSQAAPHAEIVLSNM	67
Db	522	VDSGGKGLLVKGA	VENVLERSTHIQLLDGSRELDQYSRDILQS---	LHDMLSAL 577
QY	68	RVAFDRMFLDVKEPF	-----EPSVHEAILEPPEYMFQGVIRPLVNFRESYVGNVSV	121
Db	578	R-CLGFAYSDVPSDFATYDGS	EDHPAHQQLLPNSYSSIESNLV-----	FVGFVGL 627
QY	122	FGVMEQLKQGDVKVLISN	HOEADPAVIALMLET-TNPHISENIYVAGDRVITDPLCK	180
Db	628	RDPPEKREVRQA-----	IADCR-----AGIRVMVITGDKNSTAEICREIGVFEADEDISS	678
QY	181	PFSMGRNLLCVYSKK-HMNDV	PELAEMKRSNTRSLKEMALLRGSKIIWIAPSGGRDR	239
Db	679	RSUTGKEFMDVKDQKNHLR	QTGGL--LFSRAEPKHQEIIVLLKEDGEVAMTGDGVNDA	736
QY	240	P 240		
Db	737	P 737		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 16, 2004, 02:08:21 ; Search time 102 Seconds
(without alignments)
2034.820 Million cell updates/sec

Title: US-09-926-805-7

Perfect score: 1949

Sequence: 1 GSMHGHSTRFIDARSQDL.....KGLASTPSVLSQPLQFLD 374

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DRV=xlh
-O=/cgn2_1/USPTO_spool/US09926805/runat_11032004_141821_4079/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09926805 @CNC 1.1.69 @runat_11032004_141821_4079 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/6C COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1380	70.8	1380	6	5210189-1
2	1380	70.8	1445	1	US-08-087-732-1
3	1380	70.8	1445	6	5210189-4
4	1361	69.8	1104	3	US-09-000-092-3
5	1360	69.8	1104	3	US-09-000-092-5
6	1358	69.7	1104	3	US-09-000-092-9
7	1316	67.5	1104	3	US-09-000-092-1
8	1305	67.0	1104	3	US-09-000-092-7
9	982	50.4	3397	6	5210189-3
10	536	27.5	1230025	4	US-09-198-452A-1
11	103.5	5.3	1011	4	US-09-134-001C-2712
12	100.5	5.2	10061	4	US-09-221-017B-672

13	99	5.1	1257	3	US-08-844-054-1	Sequence 1, Appli
14	99	5.1	1257	3	US-09-347-333-1	Sequence 1, Appli
15	98	5.0	4965	4	US-08-961-527-143	Sequence 143, App
16	97	5.0	1281	4	US-09-134-000C-996	Sequence 996, App
17	96.5	5.0	1728	4	US-09-328-352-3595	Sequence 3595, Ap
18	95	4.9	5583	4	US-09-312-283C-372	Sequence 372, App
19	95	4.9	10348	2	US-08-457-273B-41	Sequence 41, Appl
20	95	4.9	10348	3	US-08-556-419-13	Sequence 13, Appl
21	95	4.9	10348	3	US-09-041-886-14	Sequence 14, Appl
22	94	4.8	5672	4	US-09-023-655-1392	Sequence 1392, Ap
23	94	4.8	10366	1	US-08-246-982A-5	Sequence 5, Appli
24	94	4.8	10366	1	US-08-453-265-5	Sequence 5, Appli
25	93.5	4.8	1362	4	US-09-252-991A-14697	Sequence 14697, A
26	92.5	4.7	1888	4	US-09-270-957-27	Sequence 27, Appl
27	92.5	4.7	1830121	4	US-09-557-884-1	Sequence 1, Appli
28	92.5	4.7	1830121	4	US-09-643-990A-1	Sequence 1, Appli
29	92	4.7	4758	3	US-09-191-647-1	Sequence 1, Appli
30	92	4.7	4758	3	US-09-540-245A-1	Sequence 1, Appli
31	92	4.7	4758	3	US-09-540-153-1	Sequence 1, Appli
32	88.5	4.5	1209	4	US-09-252-991A-1340	Sequence 1340, Ap
33	88.5	4.5	1260	4	US-09-252-991A-1263	Sequence 1263, Ap
34	88.5	4.5	1572	4	US-09-252-991A-1220	Sequence 1220, Ap
35	88	4.5	1827	4	US-09-134-001C-558	Sequence 558, App
36	88	4.5	2118	4	US-09-328-352-3310	Sequence 3310, Ap
37	88	4.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
38	88	4.5	441529	3	US-09-103-840A-1	Sequence 1, Appli
39	87.5	4.5	2733	4	US-08-997-685A-1	Sequence 1, Appli
40	87.5	4.5	3025	4	US-09-976-594-552	Sequence 552, App
41	87.5	4.5	3760	4	US-09-976-594-213	Sequence 213, App
42	87.5	4.5	6765	2	US-08-677-010-2	Sequence 2, Appli
43	87.5	4.5	6765	2	US-08-790-519-2	Sequence 2, Appli
44	87.5	4.5	13086	4	US-08-956-171E-16	Sequence 16, Appl
45	87.5	4.5	80161	3	US-09-036-987A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
5210189-1
; Patent No. 5210189
; APPLICANT: MURATA, NORIO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING GLYCEROL
; 3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/587,676
; FILING DATE: 25-SEP-1990
; SEQ ID NO:1:
; LENGTH: 1380
5210189-1

Alignment Scores:
Pred. No.: 4.03e-162 Length: 1380
Score: 1380.00 Matches: 252
Percent Similarity: 84.8% Conservative: 57
Best Local Similarity: 69.2% Mismatches: 55
Query Match: 70.8% Indels: 0
DB: 6 Gaps: 0

US-09-926-805-7 (1-374) x 5210189-1 (1-1380)

QY	6	HisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGlyIleGlnArg	25
Db	280	CATTCCTCCCTACTCTTCTGGATGCGCGAAGTGAACAGATCTTTTATCTGGTATCAAGAA	339
QY	26	GluLeuGluAlaGlyThrLeuProlyshisIleAlaGlnAlaMetGluGluLeuTyrGln	45
Db	340	GAAGCTGAAGCTGGAGGTTGCCAGCAATGTTCCAGCAGGAATGGAAGATTGATTGG	399
QY	46	AsnTyrIlysAenAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSer	65
Db	400	AACTACAAAAATGCAGTCTTTAAGTAGTAGGAGCTTCCAGGCGAGATGAACACTGTGTATCA	459

```

QY 66 AsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSer 85
Db 460 AACATGTCGTCTGTTGGATCGCATCTCTGGTGTGAGGATCTTTATCTTATAT 519
QY 86 ProTyrHisGluAlaIleLeuGluProPheAspTyrTyrMetPheGlyGlnAsnTyrIle 105
Db 520 CCATATCATAAAGCAGTCAGAGAACCATTTGACTACTACATGTTTGCCATACATCATC 579
QY 106 ArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet 125
Db 580 CGTCCCTTATGATTTCAAAAATTCGTACGTGGAAATCTCTATATTTCTCTGAGCTG 639
QY 126 GluGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAla 145
Db 640 GAAGACAAAGATTCGACAGGACACAATATCGTTGTATATCAAAACCATCAAGTGAAGCT 699
QY 146 AspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIle 165
Db 700 GATCCGGCTGCTATTTCTCTATTTGCTTGAAGCAATCTCTCTTCATAGGAGAACATT 759
QY 166 IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly 185
Db 760 AATGTGTGCTGGTATCGATCATCATCTATTCCTTTTGTAAAGCGGTTCAGTATGGGA 819
QY 186 ArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGlu 205
Db 820 AGGAACCTCATATGTGTTTACTCGAAAAGACATGAATGTTGATCCTGAGCTTGTGAC 879
QY 206 MetLysIysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGly 225
Db 880 ATGAAAGAAAGAAACCAACACACGAACTTAAAGAGATGCTTCAATGCTAAGTCTGGC 939
QY 226 SerLysIleIleTrpIleAlaProSerGlyGlyArgAspArgProAspProIleThrAsn 245
Db 940 GGTCAACTTATATGATTCACCAAGCGGTGGAGGAGCGCCCGCAATCTTCTACTGGG 999
QY 246 GlnTrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAsp 265
Db 1000 GAATGTTTCTCTGACCCCTTGTATGCTTCTCGTAGACAACATGAGAAGACTGGTTGAA 1059
QY 266 HisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 285
Db 1060 CATCTCGGCTCTCGCATATATATCAATGCTTCTGTTGCTATGACATCATGCC 1119
QY 286 ProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPheHisGlyThr 305
Db 1120 CCTCCACCCAGGTTGAGAAAGAAATCGGAGAGAAAGATTAGTTGGGTTTCAAGTACT 1179
QY 306 GlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerPro 325
Db 1180 GGAATATCAATTGCTCTGAAATCACTTCTCAGACGTCACAGACGTCGAGAGCCCT 1239
QY 326 GluGluAlaLysAlaLysSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLys 345
Db 1240 AATGAGCGCAAGAGCATACAGCCAGCTTTGTACAACTCGGTGATGAACAAATACAG 1299
QY 346 ValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrProSerValSer 365
Db 1300 ATCTTAAACTCTCGGATTAAACACAGAGAGGAGTAGAGATCACTTCAAGGTCTCT 1359
QY 366 LeuSerGlnPro 369
Db 1360 TTGTCAACACCT 1371

```

RESULT 2

```

US-08-087-732-1
; Sequence 1, Application US/08087732
; Patent No. 5516667
; GENERAL INFORMATION:
; APPLICANT: Nishizawa, Osamu
; TITLE OF INVENTION: CHILLING RESISTANT PLANTS AND THEIR
; TITLE OF INVENTION: PRODUCTION
; NUMBER OF SEQUENCES: 2

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,732
; FILING DATE: 16-JUN-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP92/00024
; FILING DATE: 14-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 15883/1991
; FILING DATE: 16-JAN-1991
; APPLICATION DATA:
; APPLICATION NUMBER: JP 283807/1991
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1445 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16...1392
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 16...285
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 286...1392
; US-08-087-732-1

```

Alignment Scores:

```

Pred. No.: 4,35e-162 Length: 1445
Score: 1380.00 Matches: 252
Percent Similarity: 84.89% Conservative: 57
Best Local Similarity: 69.23% Mismatches: 55
Query Match: 70.81% Indels: 0
DB: 1 Gaps: 0

```

US-09-926-805-7 (1-374) x US-08-087-732-1 (1-1445)

```

QY 6 HisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGlyIleGlnArg 25
Db 295 CATTCCTCCGTACTTCTTGTGATGCGGAAGTGAACAAGATCTTTATCTGGTATCAAGAG 354
QY 26 GluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGluLeuTyrGln 45
Db 355 GAAGCTGAAGCTGGAAGGTTCCACGAAATGTTCCAGCAGGAATGGAAGAATTGTTATGG 414
QY 46 AsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSer 65
Db 415 AACTACAAAATATGCGATTTTAAAGTAGTGAGCTTCCAGGCGAGATGAACCTGTTGATCA 474
QY 66 AsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSer 85

```

```
Db 475 AACATGCTGTGCTTTTGAATCCCATGCTTCTTGCTGTGGAGATCCTTTATCTTTAAT 534
Qy 86 ProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIle 105
Db 535 CCATATCATAAAGCAGTCAGAACCAATTTGACTACTACATCTTTGTCATACATACATC 594
Qy 106 ArgProLeuValAspPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet 125
Db 595 CGTCTCTTATTGATTTCAAAATTCGTACGTGGAATGCTTCTATATCTCTCAGCTG 654
Qy 126 GluGluGlnLeuLysGlyAspLysValValLeuIleSerAsnHisGlnThrGluAla 145
Db 655 GAAGCAAGATTTCGACAGGACACAAATATCGTGTGATATCAAAACCATCAAAAGTGAAGCT 714
Qy 146 AspProAlaValIleAlaLeuMetLeuGluThrAsnProHisIleSerGluAsnIle 165
Db 715 GATCCGGCTGTCATTTCTATTGCTTGNAGCACATCTCTTTCATAGGAGAACAT 774
Qy 166 IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly 185
Db 775 AAATGTGTGGCTGTGATCGAGTCATCATCTGATCCTCTTTGTAAGCCGTTTCAGTATGGGA 834
Qy 186 ArgAsnLeuLeuLysValTyrSerLysLysGluMetAlaLeuLeuLeuArgGlyGly 205
Db 835 AGAACCTCATATGTTTACTCGAAAAAGCACATGAATGTGATCCTCTGAGCTTGTTCGAC 894
Qy 206 MetLysLysArgSerAsnThrArgSerLysGluMetAlaLeuLeuLeuArgGlyGly 225
Db 895 ATGAAAAAGAAAGCAACACACGAACTTAAAGAGAGTGGCTACATGCTAAAGTCTGGC 954
Qy 226 SerLysIleIleTyrPheAlaProSerGlyGlyArgAspArgProAspProIleThrAsn 245
Db 955 GGTCAACCTTATATGGATTGACCAAGCGGTGGAGGAGACCCCGCAATCTCTTCTATGGG 1014
Qy 246 GlnTyrPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAsp 265
Db 1015 GAATGTTTCTGCACCCCTTTGATGCTTCTCGGTAGACAAATGAGAACATCTGGTTGAA 1074
Qy 266 HisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 285
Db 1075 CAITCTGGCGCTCCTCGACATATATATCAATGCTCTTTGCTTGTATGACATCATGCC 1134
Qy 286 ProLeuGlnValGluLysGluIleGlyGluLysArgLeuLeuSerPheHisGlyThr 305
Db 1135 CCTCCACCCCGAGTTGAGAAAGAAATCGGAGAGAAAGATTAGTTGGGTTTTCACGGTACT 1194
```

RESULT 3

```
5210189-4
; Patent No. 5210189
; APPLICANT: MURATA, NORIO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING GLYCEROL
; 3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/587,676
; FILING DATE: 25-SEP-1990
; SEQ ID NO: 4:
```

```
5210189-4
; LENGTH: 1445
```

Alignment Scores:

```
Score: 4,35e-162 Length: 1445
Pred. No.: 1380.00 Matches: 252
Percent Similarity: 84.89% Conservative: 57
Best Local Similarity: 69.23% Mismatches: 55
Query Match: 70.81% Indels: 0
DB: 6 Gaps: 0
```

US-09-926-805-7 (1-374) x 5210189-4 (1-1445)

```
Qy 6 HisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuSerGlyIleGlnArg 25
Db 295 CATTCCTCGTCTCTTCTTGATGCGCGAAGTGAAACAAGATCTTTATCTGGTATCAAGAAG 354
Qy 26 GluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGluLeuTyrGln 45
Db 355 GAAGCTGAAGCTGGAAGGTTGCCAGCAATGTTCCAGCAGCAATGGAAGATTGTTATGG 414
Qy 46 AsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisIleGluAspIleValLeuSer 65
Db 415 AACTACAAAATGCAAGTTTAAAGTAGTGGAGCTTCCAGGGCAGATGAACTGTTGTATCA 474
Qy 66 AsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSer 85
Db 475 AACATGCTGTGTTTGTGTCATGCTTCTTGCTGTGGAGGATCCTTATACCTTTAAT 534
Qy 86 ProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIle 105
Db 535 CCATATCATAAAGCAGTCAGAGAACCATTTGACTACTACATGTTGTCATACATCATC 594
Qy 106 ArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet 125
Db 595 CGTCTCTTATTGATTTCAAAATTCGTACGTTGGAATGCTTCTATATCTCTGAGCTG 654
Qy 126 GluGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAla 145
Db 655 GAAGCAAGATTTCGACAGGACACAAATATCGTGTGATATCAAAACCATCAAAAGTGAAGCT 714
Qy 146 AspProAlaValIleAlaLeuMetLeuGluThrAsnProHisIleSerGluAsnIle 165
Db 715 GATCCGGCTGTCATTTCTATTGCTTGAAGCACAAATCTCTCTCATAGGAGAACAT 774
Qy 166 IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly 185
Db 775 AAATGTGTGGCTGTGATCGAGTCATCATCTGATCCTCTTTGTAAGCCGTTTCAGTATGGGA 834
Qy 186 ArgAsnLeuLeuLysValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGlu 205
Db 835 AGAACCTCATATGTTTACTCGAAAAAGCACATGAATGTGATCCTCTGAGCTTGTTCGAC 894
Qy 206 MetLysLysArgSerAsnThrArgSerLysGluMetAlaLeuLeuLeuArgGlyGly 225
Db 895 ATGAAAAAGAAAGCAACACACGAACTTAAAGAGAGTGGCTACATGCTAAAGTCTGGC 954
Qy 226 SerLysIleIleTyrPheAlaProSerGlyGlyArgAspArgProAspProIleThrAsn 245
Db 955 GGTCAACCTTATATGGATTGACCAAGCGGTGGAGGAGACCCCGCAATCTCTTCTATGGG 1014
Qy 246 GlnTyrPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAsp 265
Db 1015 GAATGTTTCTGCACCCCTTTGATGCTTCTCGGTAGACAAATGAGAACATCTGGTTGAA 1074
Qy 266 HisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 285
Db 1075 CAITCTGGCGCTCCTCGACATATATATCAATGCTCTTTGCTTGTATGACATCATGCC 1134
Qy 286 ProLeuGlnValGluLysGluIleGlyGluLysArgLeuLeuSerPheHisGlyThr 305
Db 1135 CCTCCACCCCGAGTTGAGAAAGAAATCGGAGAGAAAGATTAGTTGGGTTTTCACGGTACT 1194
```

QY 306 GlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerPro 325
 Db 1195 GGACTATCAATGCTCTGGAATCAACTTCTCAGAGCTCACACGACTGCGAGACCCCT 1254
 QY 326 GluGluAlaLysAlaIleTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLys 345
 Db 1255 AATGAGCGGAAAGACATACAGCCAGCTTTGTACAAGTCGGTGAATGAACAATACGAG 1314
 QY 346 ValIleHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrProSerValSer 365
 Db 1315 ATCTTAACCTCGGATTAACACAGAGGAGTAGAAGCATCAACTTCAGGGTCTCT 1374
 QY 366 LeuSerGlnPro 369
 Db 1375 TTGTCACACCT 1386

RESULT 4

US-09-000-092-3
 ; Sequence 3, Application US/09000092
 ; Patent No. 6160203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihito
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/000,092
 ; FILING DATE: 26-JAN-1998
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/JP96/01844
 ; FILING DATE: 03-JUL-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 192123/1995
 ; FILING DATE: 27-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bent, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/916
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1104 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1104
 ; US-09-000-092-3

Alignment Scores:
 Pred. No.: 6,54e-160
 Score: 1361.00
 Percent Similarity: 85.56%
 Best Local Similarity: 68.12%
 Query Match: 69.83%
 Length: 1104
 Matches: 250
 Conservative: 64
 Mismatches: 53
 Indels: 0

DB: 3 Gaps: 0
 US-09-926-805-7 (1-374) x US-09-000-092-3 (1-1104)
 QY 3 MetHisGlyHisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGly 22
 Db 1 ATGGCTAGCCACTCCGCAAAATTTCTCGATGTTGCTCTGAAGAAGAGTTGCTCTCTCTGC 60
 QY 23 IleGlnArgGluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGlu 42
 Db 61 ATCAAGAAGGAAACAGAAAGCTGCAAAAGCTGCCCTCAAAATGTTGCTGCGAGAAATGAAGAA 120
 QY 43 LeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIle 62
 Db 121 TTGTATCAGAATTATAGAAATGCTGTTATAGAGATGGAAATCCAAAGCGATGAAT 180
 QY 63 ValLeuSerAsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPhe 82
 Db 181 GTTCTGTCTAACATGACTGTTGCAATTAGATCGCATATTTGTTGGATGTGGAGGATCCTTTT 240
 QY 83 GluPheSerProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGln 102
 Db 241 GTCTTCTCATCACACCAACAAAGCAATTCGAGAGCCCTTTTGATTACTATACATTTTGGCCAG 300
 QY 103 AsnTyrIleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPhe 122
 Db 301 AACTATATACGGCCATTGATGATTTTGGAAATTCATTGCTGGTGGTAACTTCTCTTTTC 360
 QY 123 GlyValMetGluGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGln 142
 Db 361 AAGGATATAGAAGAGAGAGCTTAAGCAGGCTGACACATCATCTTAATGCCAACCATCAA 420
 QY 143 ThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSer 162
 Db 421 AGTGAAGCAGATCCCGCAGTGTGATTCATTACTCTGGAGAGACAAATTCATAATCGCA 480
 QY 163 GluAsnIleIleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPhe 182
 Db 481 GAAACTTGTATCATATAGCAGGTGATCGAGTTATTAACAGATCCTCTTTTGCAGGCCCTTT 540
 QY 183 SerMetGlyArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGlu 202
 Db 541 AGCATGGGAAGGAATCTTCTTTGTTTACTTAAGAAGCACATGTATGATGATCCCGAG 600
 QY 203 LeuAlaGluMetLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeu 222
 Db 601 CTGTGTGATGTAAGAAAGAGCAATACAGAGGTTTGAAGAGTTGGTCTTACTTTA 660
 QY 223 ArgGlyGlySerLysIleIleTrpIleAlaProSerGlyGlyArgAspArgProAspPro 242
 Db 661 AGAGTGGTTCAAAAATTAATCTGGATTGCCATCCAGTGTGGAGAGAGATCGTCCAGATGCT 720
 QY 243 IleThrAsnGlnTrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArg 262
 Db 721 GTCACTGGTGAATGGTACCCAGCACCCCTTTGATGCTTCTTCAGTGGACCAATGAGAAGG 780
 QY 263 LeuValAspHisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAsp 282
 Db 781 CTTATTCACATTCGATGTTCTCGGCATTTGTTCCCTTCTTATTTATGTCATGAC 840
 QY 283 IleMetProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPhe 302
 Db 841 ATCATGCCCTCCCTCACAGGTCGAAATTCGAAATTCGAGAAAAAGAGATGATGCTTT 900
 QY 303 HisGlyThrGlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCys 322
 Db 901 AATGGCGCGGTTTGTCTGTGCTCCTCGAAATTCAGCTTCGAGGAAATTCGCTACCCAC 960
 QY 323 GlySerProGluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGlu 342
 Db 961 AAAAATCTGAGAGGTTAGGAGGACATCTCAAGGCATCTGTTGATTCGTGCGCCATG 1020
 QY 343 GlnTyrLysValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrPro 362

Db 1021 CAATCAATGTGCTCAAAACGGCTATCTCCGCAACAGGACTAGGAGCTTCACTGCG 1080
 Qy 363 SerValSerLeuSerGlnPro 369
 Db 1081 GATGCTCTTTTGTCAACCT 1101

RESULT 5

US-09-000-092-5
 ; Sequence 5, Application US/09000092
 ; Patent No. 6160203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihiro
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/000,092
 ; FILING DATE: 26-JAN-1998
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: NO PCT/JP96/01844
 ; FILING DATE: 03-JUL-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 192123/1995
 ; FILING DATE: 27-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bent, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/916
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1104 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1104
 ; US-09-000-092-5

Alignment Scores:

Pred. No.: 8 72e-160 Length: 1104
 Score: 1360.00 Matches: 250
 Percent Similarity: 85.56% Conservative: 64
 Best Local Similarity: 68.12% Mismatches: 53
 Query Match: 69.78% Indels: 0
 DB: 3 Gaps: 0

US-09-926-805-7 (1-374) x US-09-000-092-5 (1-1104)

Qy 3 MethHisGlyHisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuSerGly 22

Db 1 ATGGCTAGCCACTCCGCAAAATTTCTCGATGTTGGCTCGAGAGAGTTCCTCTCTGCG 60

Qy 23 IleGlnArgGluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGlu 42

Db 61 ATCAAGAGGAAACAGAAAGCTGGAAGCTGCTCAAAATGTTGCTGCAAGAAATGAAGAA 120
 Qy 43 LeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIle 62
 Db 121 TTGTATCAGAAATTAGAAATGCTGTTATTGAGAGTGGAAATCCAAAGCAGATGAAT 180
 Qy 63 ValLeuSerAsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPhe 82
 Db 181 GTTCTGTCTAAACATGACTGTTGCATTAGATCGCATATTGTTGGATGTGGAGATCCTTT 240
 Qy 83 GluPheSerProTyrHisGluAlaIleLeuGlnProPheAsnTyrTyrMetPheGlyGln 102
 Db 241 GTCCTTCTCATCACCAACAGCAATTCGAGAGCTTTTGATTACTACTATCTTTTGGCCAG 300
 Qy 103 AsnTyrIleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPhe 122
 Db 301 AACTATATACGGCCATTGATTGATTTTGGAAATTCATTGCTGGTAACTCTTCTCTTTTC 360
 Qy 123 GlyValMetGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGln 142
 Db 361 AAGGATATAGAGAGAGAGCTTTAAGCAGGCTGACACATCATCTTAATGTCCAACCATCAA 420
 Qy 143 ThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSer 162
 Db 421 AGTGAAGCAGATCCGCGAGTGTGCTTCTGAGAGAGCAAAATCTAATATCGCA 480
 Qy 163 GluAsnIleIleTyrValAlaGlyAspArgValIleThrAspProLeuLysGlySerPro 182
 Db 481 GAAAACCTTGATCTACATGACAGAGTGTGCTGAGTAAACAGATCCTCTTTGCAAGCCCTT 540
 Qy 183 SerMetGlyArgAsnLeuLysValTyrSerLysLysHisMetAsnAspValProGlu 202
 Db 541 AGCATGGAGAGGAATCTCTTGTGTTTACTCTAAGAGACACATGTATGATGATCCCGAG 600
 Qy 203 LeuAlaGluMetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeu 222
 Db 601 CTCACAGAAACAAAAGAGAAAGCAACACAGAGTCTTAAGGAGAGTGGCTTACTCTTA 660
 Qy 223 ArgGlySerLysIleIleTyrPheAlaProSerGlyGlyArgAspArgProAspPro 242
 Db 661 AGAGTGGATCAACAATAATATGATGATGACCCAGTGTGAGTGGGACCGCGCGATCCC 720
 Qy 243 IleThrAsnGlnTyrPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArg 262
 Db 721 TCGACTGAGAGATGTTACCCAGCACCTTTGATGCTCTTCAGTGGAGCAACATGAGAGG 780
 Qy 263 LeuValAspHisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuLysHisAsp 282
 Db 781 CTTATTCAACATTCGGATGTTCTCGGCAATTTGTTTCCCTTGTCTTATTATGTCATGAC 840
 Qy 283 IleMetProProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPhe 302
 Db 841 ATCATGCCCTCCCTCCACAGGTGCAAAATTTGAAATTCAGTTCGAGGAAATGTGTACCAC 900
 Qy 303 HisGlyThrGlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCys 322
 Db 901 AATGGGCGGGTGTCTGTGGCTCTGAAATTCAGTTCGAGGAAATGTGTACCAC 960
 Qy 323 GlySerProGluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGlu 342
 Db 961 AAAAACTCTGAGGAGGTTAGGAGGCATCTCAAGGACATGTTGTTGATTCTGTGGCCATG 1020
 Qy 343 GlnTyrLysValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrPro 362
 Db 1021 CAATACAATGTCTCAAAACGGCTATCTCCGCAAAACAGAGGACTAGGAGCTTCAACTGCG 1080
 Qy 363 SerValSerLeuSerGlnPro 369
 Db 1081 GATGCTCTTTTGTCAACCT 1101

RESULT 6

US-09-000-092-9

```

; Sequence 9, Application US/09000092
; Patent No. 6160203
; GENERAL INFORMATION:
; APPLICANT: FERRI, Stefano
; APPLICANT: TOGURI, Toshihiro
; TITLE OF INVENTION: DNA STRANDS CODING FOR
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,092
; FILING DATE: 26-JAN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP96/01844
; FILING DATE: 03-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 192123/1995
; FILING DATE: 27-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16987/916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1104
; US-09-000-092-9

Alignment Scores:
Pred. No.: 1.55e-159 Length: 1104
Score: 1358.00 Matches: 249
Percent Similarity: 85.99% Conservative: 64
Best Local Similarity: 68.41% Mismatches: 51
Query Match: 69.68% Indels: 0
DB: Gaps: 0

US-09-926-805-7 (1-374) x US-09-000-092-9 (1-1104)

QY 6 HisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGlyIleGlnArg 25
Db 10 CACTCCCGCAAAATTTCTGATGTTGCTGAGAGAGAGTTGCTCTCTCGCATCAAGAAG 69
QY 26 GluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGluLeuTyrGln 45
Db 70 GAAACAGAGCTGGAAGTGTCTCCAAATGTGTGAGGAGTGAAGATGATGATG 129
QY 46 AsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSer 65
Db 130 AATTATAGAAATGCTGTTATTGAGAGTGGAAATCCAAAGGCAGATGAAATTTCTGTCT 189
QY 66 AsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSer 85

```

```

Db 190 AACATGACTGTTGCATTAGATCGCATATATGTTGGATGTGGAGGATCCCTTTTGTCTCTCA 249
QY 86 ProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIle 105
Db 250 TCACACCACAAAGCAATTCGAGAGCCTTTTGTATTACTACATTTTGGCCAGAACTATATA 309
QY 106 ArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet 125
Db 310 CGGCATTGATTGATTTGGAAATTCATGTTGGTAACCTTTCTCTTTCAAGATATA 369
QY 126 GluGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAla 145
Db 370 GAAGAGAACTTAAGCAGGGTGACAAACATCATCTTATGTCCACCATCAAGGAACA 429
QY 146 AspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIle 165
Db 430 GATCCCGCAGTGATTGCAATTACTTCTGGAAGAAGCAAAATTCACATAATCGCAAAACATTG 489
QY 166 IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly 185
Db 490 ATCTACATAGCAGGTGATCGAGTTATAACAGATCCTCTTTGCAAGCCCTTTAGCATGGA 549
QY 186 ArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGlu 205
Db 550 AGGAATCTTCTTTGTGTTTACTCTAAGAAGACACATGTATGATGATGCCGAGCTTGTGAT 609
QY 206 MetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGly 225
Db 610 GTAAGAAAGAGCAAAATACAGAGTTTGAAGAGTTGGTCTTACTTTTAAAGAGGTGGT 669
QY 226 SerLysIleIleTyrPheAlaProSerGlyGlyArgAspArgProAspProIleThrAsn 245
Db 670 TCAAAATAATCTGGAATTCACCCAGTGTGGAAGAGATCTCCAGATGCTGCTACTGGT 729
QY 246 GlnTrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAsp 265
Db 730 GAATGGTACCACGACCCCTTTTGATGCTTCTTCAGTGGACACATGAGAAAGCTTATTCAA 789
QY 266 HisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 285
Db 790 CATTGGATGTTCTGGCAATTTGTTCCCTTCTTATTTATGTCATGACATCATGCC 849
QY 286 ProProLeuGlnValGluLysGluIleGlyGluLysArgLeuLeuSerPheHisGlyThr 305
Db 850 CTTCCCTCACAGGTCGAAATTTGAATTTGGAGAAAAAGAGTGATTCCTTTAATGGGCG 909
QY 306 GlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerPro 325
Db 910 GGTGTTGTGTGGCTCTCTGAAATCAGCTTCGAGGAAATTTGCTGTACCCACCAAAATCT 969
QY 326 GluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLys 345
Db 970 GAGGAGTTAGGAGGAGCATCTCAAGCAGCTGTTGATTTCTGCGCATCAATACAT 1029
QY 346 ValLeuHisSerAlaValHisGlyGlyGlyGlyLeuGluAlaSerThrProSerValSer 365
Db 1030 GTGCTCAAAACGGCTATCTCCGCAACCAAGAGCTTAGAGCTTCAACTCGGAGTCTCT 1089
QY 366 LeuSerGlnPro 369
Db 1090 TTGTCACAACT 1101

RESULT 7
US-09-000-092-1
; Sequence 1, Application US/09000092
; Patent No. 6160203
; GENERAL INFORMATION:
; APPLICANT: FERRI, Stefano
; APPLICANT: TOGURI, Toshihiro
; TITLE OF INVENTION: DNA STRANDS CODING FOR
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: FOLEY & LARDNER
 STREET: 3000 K Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/000,092
 FILING DATE: 26-JAN-1998
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/JP96/01844
 FILING DATE: 03-JUL-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 192123/1995
 FILING DATE: 27-JUL-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16887/916
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1104 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1104
 US-09-000-092-1

Alignment Scores:

Pred. No.: 2,72e-154 Length: 1104
 Score: 1316.00 Matches: 243
 Percent Similarity: 93.11% Conservative: 62
 Best Local Similarity: 66.21% Mismatches: 62
 Query Match: 67.52% Indels: 0
 DB: 3 Gaps: 0

US-09-926-805-7 (1-374) x US-09-000-092-1 (1-1104)

QY 3 MethHisGlyHisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuSerGly 22
 Db 1 ATGGCTAGCCACTCTCGACATTATCGTAACGTTGTTCTGCAGAGAGCTGATATCTGAA 60
 QY 23 IleGlnArgGluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGlu 42
 Db 61 ATAAAAGGGAATCAGAAATGGAAGTTACCTAAAGCTGTGCTATGCTATGGAGGA 120
 QY 43 LeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIle 62
 Db 121 CTTTTCCTACTACTCGCAATGTCAGTCTCTTTCAAGTGAATTTCTCATGCTGATGAATA 180
 QY 63 ValLeuSerAsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPhe 82
 Db 181 GTGTGTCAACATGAGTGTATGCTGATTTGTTGTTGATATTGGAGATTTGAGGACCTTTT 240
 QY 83 GluPheSerProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGln 102
 Db 241 GTATTTCACCGTTTCACAAAGCTATTTCGAGAGCTGCTGACTATTATTCCTTTGGTCAA 300
 QY 103 AsnTyrIleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPhe 122
 Db 301 GATTACATTGGCCATTGGTAGATTGGAATTCATATGTTGGTAACATCGCCATTTC 360

QY 123 GlyValMetGluGluGlnLeuLysGlnGlyAspYsValValLeuIleSerAsnHisGln 142
 Db 361 CAAGAAATGGAGGAAGCTTAAGCAGGGTGACACATCATCTTAATGTCCAAACCATCAA 420
 QY 143 ThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSer 162
 Db 421 AGTGAAGCAGATCCCGCAGTGTATTCCTTGGAGAGACAAATTCATAATCGCA 480
 QY 163 GluAsnIleIleTyrValAlaGlyAspArgValIleThrAspProLeuLysCysIleSer 182
 Db 481 GAAACTTGATCTACATGACAGTGTATTCAGTATTAACAGATCCCTTTTGCAGCCCTTT 540
 QY 183 SerMetGlyArgAsnLeuLysCysValTyrSerLysLysHisMetAsnAspValProGlu 202
 Db 541 AGCATGGGAAGGAATCTTTTGTGTTTACTCTAAGAAGCACATGATGATGATCCCCGAG 600
 QY 203 LeuAlaGluMetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeu 222
 Db 601 CTTTGTGATGTAAAGAAAGAGCAATAACAAGGAGTTTGAAGAGTGTGCTTACTTTTA 660
 QY 223 ArgGlyGlySerLysIleIleTyrIleAlaProSerGlyGlyArgAspArgProAspPro 242
 Db 661 AGAGTGGTTCAAAAATAATCTGGATTGCCACCCAGTGTGGAAGAGATCGTCCAGATGCT 720
 QY 243 IleThrAsnGlnTyrPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArg 262
 Db 721 GTCACTGGTGAATGGTATCCACAGCACCTTTGATGCTTCTTCAGTGGACACATGAGAAGG 780
 QY 263 LeuValAspHisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAsp 282
 Db 781 CTTATTCAACATTGGATGTTCTGGGCAATTTGTTCCCTTGTCTTATTAATGTCATGAC 840
 QY 283 IleMetProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPhe 302
 Db 841 ATCATGCCCTCCCTCACAGTTCGAATTTGAATTTGGAGAAAAGAGTGAATGGCTTT 900
 QY 303 HisGlyThrGlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCys 322
 Db 901 AATGGGGGGGTTTGTCTGTGGCTTCCTGAAATCAGTTCGAGGAAATTTGCTTACCCAC 960
 QY 323 GlySerProGluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGlu 342
 Db 961 AAAAATCTGAGGAGTTAGGAGGCATCTCAAGGACCTGTTTGAATCTGTGGCCATG 1020
 QY 343 GlnTyrLysValLeuHisSerAlaValIleGlyGlyLysGlyLeuGluAlaSerThrPro 362
 Db 1021 CAATACAAATGTGCTCAAAACGGCTATCTCCGGCAACAAGGACTAGGAGCTTCAACTCG 1080
 QY 363 SerValSerLeuSerGlnPro 369
 Db 1081 GATGCTCTCTTTGTCAACACCT 1101

RESULT 8

US-09-000-092-7
 ; Sequence 7, Application US/09000092
 ; Patent No. 6160203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihiro
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/000,092
 FILING DATE: 26-JAN-1998
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/JP96/01844
 FILING DATE: 03-JUL-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 192123/1995
 FILING DATE: 27-JUL-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16887/916
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1104 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1104
 US-09-000-092-7

Alignment Scores:
 Pred. No.: 6,43e-153 Length: 1104
 Score: 1305.00 Matches: 241
 Percent Similarity: 83.47% Conservative: 62
 Best Local Similarity: 66.39% Mismatches: 60
 Query Match: 66.96% Indels: 0
 DB: 3 Gaps: 0

US-09-926-805-7 (1-374) x US-09-000-092-7 (1-1104)

```

QY 7 SerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGlyIleGlnArgGlu 26
Db 13 TCTCGCACTATCGTAACGTTGCTTCTCGAGAGAGCTGATATCTGAATAAAGGGAA 72
QY 27 LeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsn 46
Db 73 TCAGAAATGGAGGTTACTTAAAGTGTGCTTATGCTATGGAGGACTTTTCACTAC 132
QY 47 TyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsn 66
Db 133 TATCGCAATGCGAGTCTTTCAAGTGAATTTCTCATGCTGATGAATAGTGTGTCAAC 192
QY 67 MetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSerPro 86
Db 193 ATGAGTGTTATGCTGATTTGTTTGTGGATATGAGACCTTTGTATTTCCACCG 252
QY 87 TyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIleArg 106
Db 253 TTTCAAAAGCTATTCGAGAGCGCTGCTACTATTATTCCTTTGGTCAAGATTACATTGG 312
QY 107 ProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMetGlu 126
Db 313 CCAATGGTGAATTTGGAAATTCATATGTTGTAACATCGCCATTTTCCAAAGAAATGGAG 372
QY 127 GluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAlaAsp 146
Db 373 GAGAAGCTTAACGAGGCTGACACATCATCTTAATGTCACCATCAAGTGAAGAGAT 432
QY 147 ProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIleIle 166
Db 433 CCGCGAGTGTGCTTACTTCTGAGAGAGACAAATTCATTAATTCATTAATTCGAGAAACTTGATC 492
  
```

```

QY 167 TyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGlyArg 186
Db 493 TACATAGCAGGTGATCGAGTTATAACAGATCTCTTTGCAAGCCCTTTAGCATGGGAAG 552
QY 187 AsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGluMet 206
Db 553 AATCTTCTTTGGTTTACTCTAAGAAGCACATGTATGATGATCCCGAGCTTGTGTATGTA 612
QY 207 LysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGlySer 226
Db 613 AAGAAAGAGACAAATACAGAGGATTTGAAAGAGTTTGGTCTTACTTTTAAAGAGGTGTTCA 672
QY 227 LysIleIleTyrIleAlaProSerGlyGlyArgAspArgProAspProIleThrAsnGln 246
Db 673 AAAATATCTGGATTCACCCAGTGTGGAAGAGATCGTCCAGATCTGTCACTGTGTGAA 732
QY 247 TrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAspHis 266
Db 733 TGGTACCCAGCACCTTTGATGCTTCTTCAGTGGACAAACATGAGAAGGCTTATTCACAT 792
QY 267 AlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 286
Db 793 TCGAGTGTTCCTGGGCAATTTGTTCCCTTGTCTTTATATGTCATGACATCATGCCCT 852
QY 287 ProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPheHisGlyThrGly 306
Db 853 CCTCACAGTTCGAATTTGAATTTGAGAAAAGAGTGTGCTTTTATGGGGCGGT 912
QY 307 IleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerProGlu 326
Db 913 TTGCTGTGCTCTCGAAATTCAGCTTCGAGGAAATTTGCTGTACCACAAAAATCTGTAG 972
QY 327 GluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLysVal 346
Db 973 GAGGTAGGAGGCATACCTCAAGGCACTGTGTGATTTGTGGCCATGCAATACAAATGTG 1032
QY 347 LeuHisSerAlaValHisGlyLysGlyLeuGluAlaSerThrProSerValSerLeu 366
Db 1033 CTCAAACGGCTATCTCCGGCAACAGGACTTAGGAGCTTCACTGCGGATGCTCTTTG 1092
QY 367 SerGlnPro 369
Db 1093 TCACAACCT 1101
  
```

RESULT 9

5210189-3
 Patent No. 5210189
 APPLICANT: MURATA, NORIO
 TITLE OF INVENTION: DNA SEQUENCE ENCODING GLYCEROL
 3-PHOSPHATE ACYLTRANSFERASE
 NUMBER OF SEQUENCES: 4
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/587,676
 FILING DATE: 25-SEP-1990
 SEQ ID NO: 3
 LENGTH: 3397
 5210189-3

Alignment Scores:
 Pred. No.: 8,95e-112 Length: 3397
 Score: 982.00 Matches: 243
 Percent Similarity: 43.54% Conservative: 57
 Best Local Similarity: 35.27% Mismatches: 52
 Query Match: 50.38% Indels: 340
 DB: 6 Gaps: 11

US-09-926-805-7 (1-374) x 5210189-3 (1-3397)

```

QY 18 AspLeuLeuSerGlyIleGlnArgGluLeuGluAlaGlyThrLeuProLysHisIleAla 37
Db 1122 GATCTTTATCTGTTATCAAGAGAGAGCTGAAGCTGGAAGTTGCCAGCAATGTTGCA 1181
QY 38 GlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSer----- 54
  
```

Db	1182	GCAGGAATGAAGAATAATTGTATTTGGAACTACAAAAATCAGTAATGAATCCTTTTATCAC	1241
Qy	54	-----	54
Db	1242	TCACTAGTTGTAGCATTTAGTGTCTATGCTTTGSAATTAATAGATTAAAGAACAATCAGT	1301
Qy	55	-----	AlaAla 56
Db	1302	CCACTGAACCTTTGGCATTGTTGTTACTATGATGGTTTTTAGGTTTTTAAGTAGTGGAGCTT	1361
Qy	57	ProHisAlaGluAspIleValLeuSerAsnMetArgValAlaPheAspArgMet-PheLe	76
Db	1362	CCAGG-CGAGATGAACACTGTTGTATCAACATGCTGTTGCTTTTGATCGATGCTTCTT	1420
Qy	76	u-----	AspVal-- 78
Db	1421	GGTGTGGAGGTACTTCTCTGTCAATCTCTTGGATGAATTTTCTACTAGAGATGTCAT	1480
Qy	79	-----	lysGluProPheGluPh 84
Db	1481	CTAAAAATCTTTTACTTGAGCTTCCTTGGCTGTGAATATTTCCAGCATCTTATATCTTT	1540
Qy	84	eSerProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyr	104
Db	1541	TAATCCATATCATAAAGCAGTCAGAGAACCATTTCAGTACTACATGTTTGTCCATACATA	1600
Qy	104	rIleArgProLeuValAsnPhe-----	111
Db	1601	CATCCGTCCCTTATTGCA-TTTCAAGTAAGCTGGAATATGCATCATAGTTATCAGATTT	1659
Qy	111	-----	111
Db	1660	TATTCATAACTGAATGTAAATTAGAAAGCAATATCGTTTGGTTAGTATCATCATGATCTTCCA	1719
Qy	112	-----	ArgGI 113
Db	1720	GTTTTCTTCTGGCATCAAGTGTTCATCTACTAACAAACCTTGAATATGTTATGCAGAAA	1779
Qy	113	uSerTyrValGlyAsnValSerValPheGlyValMetGluGluGlnLeuLysGln----	131
Db	1780	TTCTGATCGTTGGAATGCTTCTATTTCTCTGAGCTGGAGACAGATTCGACAGGTCAT	1839
Qy	131	-----	131
Db	1840	CTTCCCTCTTCTGTGCTCTAGTTAATGCTGTTTGTCTTTCATCAACAGTTAAAGTTCCT	1899
Qy	132	-----	GlyAspLys-ValValLeuLies 139
Db	1900	TCCATTCCAGGATTTTAATCTCAACTGTTTAAATTACAGGGACACAATATCGTGTGTAT	1959
Qy	139	erAsnHisGlnThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnP	159
Db	1960	CAAAACCATCAAGTGAAGCTGATCCGGCTGTCAATTCCTATGCTTGAAGCACATCTC	2019
Qy	159	roHisIleSerGluAsnIleIle-----	166
Db	2020	CTTTCATAGGAGAGAACATTGTGAGCCTTCGAGCCTTGTCTGAGCTATTATACGCTATT	2079
Qy	167	-----	TyrValAlaGlyAspArgValIleT 175
Db	2080	GTCTCTGTTCTTATTGTTTCTTTGTTGGTTCAGAAATGTGTGGCTGTGATCGAGTCATCA	2139
Qy	175	hrAspProLeuCysLysProPheSerMetGly-----	185
Db	2140	CTGATCTCTTTGTGACCGCTTCAGTATGGGAAGGTATCAGAGCTTTCATTTCGAAATGG	2199
Qy	185	-----	185
Db	2200	TATGCTACCTGATGGCTGAAGTGAATCTTTAGCACCTTGTATGACTTTCGGGATTTCTA	2259
Qy	186	-----	ArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGluL 203

D	b	2260	CTGCTTTTCAGGAACCTCATATGTGTTTTACTCGAAAAGCACATGAATGTTTGATCCTCGAGC	2319
Q	y	203	euaLaGIuMetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeu-	222
D	b	2320	TTCGTTGCATCGAAAGAAGCAACAACAGAGCTTAAGAGGAGATGGCTACATGCTAA	2379
Q	y	222	-----	222
D	b	2380	GGTTAATGGAATAAATAGCGAGTGTCTTATTATTGATCTCATAGGAACAGAAATTAAGG	2439
Q	y	223	-----ArgGlyGlySerLysIleI	229
D	b	2440	AAACAATTGATGTTTAGTCGTCTAATGTTGATGAACCCCTAGCTGGCGGTCAACTTA	2499
Q	y	229	leTrpIleAlaProSerGlyGlyArgAspArgProAspProlIeThrAsnGlnTrpPhep	249
D	b	2500	TATGGATTGCACCAAGCGGTGGAGGACC GCCCGGAATCCTTCTACTGGGGATGGTTTC	2559
Q	y	249	ro	249
D	b	2560	CTGTAACTTATTGATGGAATACAGAGATCTTATCTGAGTCTGCTAGATATGCAAGTACT	2619
Q	y	250	-----AlaProPheAspAlaThrSerLeuAspAsnMetArgArgIeuV	264
D	b	2620	AATCTAGTAANGTTCAGGCCCTTTGATGCTTCTTCGGTAGCAACATGAGAAGACTGG	2679
Q	y	264	alaSpHisAlaGlyLeuValcylHisIleTyrrProLeuAlaIleLeuCysHisAspilem	284
D	b	2680	TTGNACATTTGGCGCTCCTGGACATATATTCNATGCTTTGCTTGTATGACATCA	2739
Q	y	284	etProProProLeu-	288
D	b	2740	TGCCCCCTCCACC-CCAGGTATTGATTGTTTTCTCTCTGCTCTCTCTCCGTTTTG	2798
Q	y	289	-----GlnValGlul	292
D	b	2799	GGCTAACCTCGGTGACTTTTCCTCATCATCACTTCATTTGTCATCTGCCAGGTTGAGA	2858
Q	y	292	ysGIuileGlyGIuLysArgLeuIleSerPheHisGlyThrGlyIleSerValalaproG	312
D	b	2859	AAGNAATCGAGAGAGAAAGATTAGTTGGGTTTTCACGGTACTGCACTATCAATTGCTCTG	2918
Q	y	312	luleAsnpheGlnGluValThrAlaSerCysGlySerProGluGlu-	327
D	b	2919	AAATCAACTTCTCAGACGCTCACAGACACTGCGAGAGCCCATAATGAGTCTGCTGTAAAC	2978
Q	y	327	-----	327
D	b	2979	ACTCAGCACCAATGACTTTAGTTTCAAGATCGAGTAATTCATCATGCAATAAATAFACA	3038
Q	y	328	-----AlaLysAlaAlaTyrrSerGlnAlaLeuTyrrAspSerValCysG	342
D	b	3039	TTCTTGTGTTATGGCAGGCGAAGAGCATACAGCCAGCTTTGTACAAGTCGGTGAATG	3098
Q	y	342	luGlnTyrrLysValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrp	362
D	b	3099	AACAATACGAGATCTTAAACTCTGCGATTAAACACAGAAGAGGAGTAGAAGCATCAACT	3158
Q	y	362	roservAlserLeuSerGlnPro	369
D	b	3159	CAAGGGTCTCTTTGTCAACAACCT	3181
RESULT 10				
US-09-198-452A-1/C				
; Sequence I, Application US/09198452A				
; Patent No. 6559294				
; GENERAL INFORMATION:				
; APPLICANT: Griffiths, R.				
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides,				
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis				
; TITLE OF INVENTION: and treatment of infection				
; FILE REFERENCE: 9710-003-999				
; CURRENT APPLICATION NUMBER: US/09/198,452A				


```

; LOCATION: (690001)..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (705001)..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (720001)..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (735001)..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (750001)..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (765001)..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (780001)..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (795001)..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (810001)..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (825001)..(840000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (840001)..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature

```

US-09-926-805-7 (1-374) x US-09-198-452A-1 (1-1230025)

QY	32	LeuProGlyHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsnTyrIysAsnAlaVal	51
Db	1104887	CTCCAGAGCCTTTATATCAGAAATTTCTGTTTTTCATCAGAATTCATCGATGCAGCT	1104828
QY	52	Leu--GlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsnMetargValAla	70
Db	1104827	ACAAAANAAGCTCGAGTCATCAAGCTGAGTGCTGTATGTTTACAATGGGTCAAGGTTATC	1104768
QY	71	PheAspArgMetPheLeuAspValIysGluProPheGluPheSerProTyrHisGluAla	90
Db	1104767	ATCGAG-----GACTTGAAAAACCCCTTCATTTCCCTCGGTATCATATAAAAA	1104720
QY	91	IleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIleArgProLeuValAsn	110
Db	1104719	ATTGAGAGCTCCTATAGATCATTTTCGTTGAGTATAGATTTTTTTTCTCTAGCTATTGAC	1104660
QY	111	PheArgGluSerTyrValGlyAsnValSerValPheGlyValMetGluGluGlnLeuIys	130

Db 1104659 GATAAAAACTCTCGAATTTTAAACCTTAAAGCAACCGCTTAAAGAAATAGAGAATACATTGCT 11046600
 Qy 131 GlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAlaProAlaValIle 150
 Db 1104599 AGGGAGATAAATGTGTGTCTCTTGGCAATTCACACAGAGAATGCCATCCCAACCTCATG 11045940
 Qy 151 AlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIleIleTyrValAlaGly 170
 Db 1104539 TATTATGCCCTTAGCGAAGACGCATCTCTGAGTTAATGGAGAACAATGTTTTGTAGCTGGA 1104480
 Qy 171 AspArgValIleThrAspProLeuCysLysProPheSerMetGlyArgAsnLeuLeuCys 190
 Db 1104479 GACCGAGTCACTTCTGATCCCTTAGCAGCACCACTTTAGCATGGGATGTGATTTATTATGT 1104420
 Qy 191 ValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGluMetLysLysArgSer 210
 Db 1104419 ATTTATTCTAAGGTCATATTGCCATCTCCACGAACCTCGGGAGAAAGCTTCTTCAT 1104360
 Qy 211 AsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGlySerLysIleIleTrp 230
 Db 1104359 AATCAGAAAAGCATGCAAAATATTAAAGACCTTGTTAAATGAAGAGGCAAAATTTATTATTC 1104300
 Qy 231 IleAlaProSerGlyGlyArgAspArgProAspProIleThrAsnGlnTrpPheProAla 250
 Db 1104299 GTAGCCCTTCGAGGAGGTCTGTACAGAAAAAAT---GCAGAGAGAAAGATTATATCTTCA 1104240
 Qy 251 ProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAspHisAlaGlyLeuVal 270
 Db 1104242 GAATTCCTCCAGAAAGCATGAAAGTGTTCGTTTATTAGCTTAAGCTTCTTAATCAAACG 1104183
 Qy 271 GlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetProProLeuGlnVal 290
 Db 1104182 ACACATTTTATCCCTTGTCTTAAAAAAGCTATGACATCTCCGCCCTCCTCCGAAATA 1104123
 Qy 291 GluLysGluIleGlyGlyLysArgLeuIleSerPheHisGlyThrGlyIleSerValAla 310
 Db 1104122 GAGAATGCAATTTGGGAGAGCGAGCCATTTTCTTT-----GCT 1104084
 Qy 311 ProGlu---IleAsnPheGlnGluValThrAlaSerCysGlySerProGluGluAlaLys 329
 Db 1104083 CCTGTTTCTTCAACTTCGA----- 1104063
 Qy 330 AlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLysValLeuHisSer 349
 Db 1104062 -----GCAGAACTATTTTTGTATGATTTATGTTCAAAGAGGAACCTTATTCAATGC 1104012
 Qy 350 AlaValHisGlyGlyLysGlyLeuGluAla 359
 Db 1104011 GATAAACACGCCCAAGAACATTAGAGCA 1103982
 RESULT 11
 US-09-134-001C-2712
 ; Sequence 2712, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Boucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; PRIORITY FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 2712
 ; LENGTH: 1011
 ; TYPE: DNA
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-2712

```
Alignment Scores:
Pred. No.: 0.0058 Length: 1011
Score: 103.50 Matches: 49
Percent Similarity: 43.08% Conservative: 35
Best Local Similarity: 25.13% Mismatches: 78
Query Match: 5.31% Indels: 33
DB: 4 Gaps: 9

US-09-926-805-7 (1-374) x US-09-134-001C-2712 (1-1011)
Qy 49 AsnAlaValLeuGlnSerAla---AlaProHisAlaGlu---AspIleValLeuSerAsn 66
Db 409 AACCGGTGATGGAAGACGGGAAATGTCATCGAGAACTAATATACCTAAACACAT 468
Qy 67 MetArgValAlaPheAspArgMetPheLeuAspVallyysGluProPheGluPheSerPro 86
Db 469 AAAAAAATA-----GATCGTGTGTTTTA-----GAACCAAGTGATGTTGAACCA 513
Qy 87 TyrHisGluAlaIle-----LeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyr 104
Db 514 ATGATGAGCGATAGAGCTTTAGACACAGCATTAATGTCITTAGACACAGGT--- 570
Qy 105 IleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyVal 124
Db 571 -----TCATTATATACAAAGTGTATATCAAAATTTATGTCCAAAGGTATT 615
Qy 125 MetGluGlnLeuLysGlnGlyAspLysValValLeuLysSerAsnHisGlnThrGlu 144
Db 616 TCAGAAGCATTTATACGTACATCTGCTCCAAACTTTATGATCTAATGTTATGACACAA 675
Qy 145 AlaAspProAlaValIleAlaLeuMetLeuGluThrAsnProHisIleSerGluAsn 164
Db 676 CCAGGC-----GAGACTGATAATTTATGATGTCMAAGACAT 711
Qy 165 IleIleTyrVal-----AlaGlyAspArgValIleThrAspProLeuCysLysPro 181
Db 712 ATTGATGCACCTTACTCGACAAAGTTGGTGACCACTTTATGATTTTGTATATGATGCTCA 771
Qy 182 PheSerMetGlyArgAsnLeuLysValTyrSerLysLysHisMetAsnAspValPro 201
Db 772 GAATCCTATAGTAAGAGTTTTCACAGGATATGAAGAAAGAAATTCGAACACAGTAGCA 831
Qy 202 GluLeuAlaGluMetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeu 221
Db 832 GTACATAAAGAACAAATTAAGAGATAGTGAATTTAGAGTTTAAACGGCATCTTAATTAGTT 891
Qy 222 -----LeuArgGlyGlySerLysIlelle 229
Db 892 GAAATATCTAATGAACACTATGTGACAGACATAACACAAAGTATTA 936

RESULT 12
US-09-221-017B-672/c
; Sequence 672, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
```



```

Db 2364 TTCCTTCATCCGTCATTCACCTCCGCAAGGCTTTT-----GATCTTGCCTTCTTT--- 2314
QY 124 ValMetGluGluGlnLeuLysGlnGlyAspLysValVal-----LeuLysSerAsn 140
Db 2313 -----GATCAAGAGGCAAGCGGTGGATCTGTGTGTGGAGCGAT 2272
QY 141 HisGlnThrGluAlaAspProAlaValIleAlaLeuMet-----LeuGluThrThr 157
Db 2271 CATGCCACAGATGTCACCCAGCTGCGCTCCCTCATTTGTGAGTCGGCTTGAGAGCACT 2212
QY 158 AsnProHisIle-----SerGluAsnIleIleTyrValAlaGlyAspArgValIle 174
Db 2211 CAG-----CACTCTCTCATCCGTACGCGCAAGCAACAATCTTCAAGCGCTTTCATGATGAATT 2155
QY 175 ThrAspProLeuCys-----LysProPheSerMetGlyArgAsn 187
Db 2154 GCGGAAACTGTTTTCAGCCCATCTCATAGAACATCAGCCCTTGTCCCAAGCGCAGAGA 2095
QY 188 LeuLeuCysValTyrSerLysLysHis 196
Db 2094 AC-GTATGCGCATATCATGATAGACAT 2069

RESULT 13
US-08-844-054-1
; Sequence 1, Application US/08844054
; Patent No. 6071892
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6071892el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,054
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9608001.5
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31454
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-844-054-1

Alignment Scores:
Pred. No.: 0.0304
Score: 99.00
Percent Similarity: 42.93%
Best Local Similarity: 22.83%
Query Match: 5.08%
Length: 1257
Matches: 42
Conservative: 37
Mismatch: 61
Indels: 45

```

```

DB: 3 Gaps: 10
US-09-926-805-7 (1-374) x US-08-844-054-1 (1-1257)
QY 15 SerGluGlnAspLeuLeuSerGlyIle-----GlnArgGluLeuGluAlaGlyThrLeu 32
Db 665 TCACAGATGCAACTGGTAAAAAATTTGTAATCAGAAAGAAATCCGCTCGCTCAATC 724
QY 33 ProLysHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaValLeu 52
Db 725 CGGAAAAGACATCTTCCATAC-----GAAATGTACCAATTTCTGGATGAAC----- 768
QY 53 GlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsnMetArgValAlaAlaPheAsp 72
Db 769 -----GTGATGGAGCGCTGACGCTGTTCGCTTCTTG 798
QY 73 ArgMetPheLeu-----AspValLysGluProPheGluPhe 84
Db 799 AAAATCTTTACTTCTTCTGTCATCTTGATGAGATTGAAGATATTGTAACAATTTGAAGCA 858
QY 85 SerProTyr-----HisGluAlaIleLeuGluProPheAsnTyrTyrMetPhe 100
Db 859 GCGCCACAGCAAGCTTGGCTCAAAAAGTCTGGCTCGTGAAGTTGTACACTTCTTTCAC 918
QY 101 GlyGlnAsnTyrIleArgProLeuValAsnPheArgGluSer---TyrValGlyAsnVal 119
Db 919 GGAGAGAGAGCGCTACAAAGAGCACTTAACATCATCTGAGCAACTCTTTGCAGGAACATC 978
QY 120 SerValPheGlyValMetGluGlnLeuLysGlnGlyAspLysValValLeuLeuSer 139
Db 979 AAAAACCTTCTGTCAAAGAG-----CTCAACAAGGACTTCGTGTGTG-----CCA 1026
QY 140 AsnHisGlnThrGluAlaAspPro-----AlaValIleAlaLeuMetLeuGluThr--- 156
Db 1027 AACTACCAAGTACAGGCAGAGCAAAACAAATATCGTGAACCTGCTCTCATCTGCT 1086
QY 157 -----ThrAsnProHisIleSerGluAsnIleIleTyrValAla 169
Db 1087 ATAGTTAACTCAAAACGCCAGCCGTCGAAGACGTCCTCAAAACGGAGCCATCTACGTAAAC 1145
QY 170 GlyAspArgVal 173
Db 1147 GCGGACCGCATC 1158

RESULT 14
US-09-347-333-1
; Sequence 1, Application US/09347333
; Patent No. 6284489
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6284489el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/347,333
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,054
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31454
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-347-333-1

Alignment Scores:
Pred. No.: 0.0304 Length: 1257
Score: 99.00 Matches: 42
Percent Similarity: 42.93% Conservative: 37
Best Local Similarity: 22.83% Mismatches: 61
Query Match: 5.08% Indels: 45
DB: 3 Gaps: 10

US-09-926-805-7 (1-374) x US-09-347-333-1 (1-1257)

Qy	15	SerGluGlnAspLeuLeuSerGlyIle-----GlnArgGluLeuGluAlaGlyThrLeu	32
Db	665	TCACAGATGCAACTGGTAAATAATGTAATCAGAAGAAATCCGCTCTGGCTCAATC	724
Qy	33	ProLysHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaValLeu	52
Db	725	CCGAAAAAGACTTCTCCATPAC-----GAAATGTACCAATTTCTGGATGAAC	768
Qy	53	GlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsnMetArgValAlaPheAsp	72
Db	769	-----GTGATGGAAGCTGACGCTGTTCGCTCTCTTG	798
Qy	73	ArgMetPheLeu-----AspValLysGluProPheGluPhe	84
Db	799	AAATCTTTACTTTCTTGCTACTTGATGAGATTGAAGATATCGTAACAATTTGAAGCA	858
Qy	85	SerProTyr-----HisGluAlaIleLeuGluProPheAsnTyrTyrMetPhe	100
Db	859	GCGCCACACGAACGCTTGGCTCAAAAAGTCTTGCTCGTGAAGTGTGTACACTTGTTCAC	918
Qy	101	GlyGlnAsnTyrIleArgProLeuValAsnPheArgGluSer---TyrValGlyAsnVal	119
Db	919	GGAGAAGAGCCCTACAAGAGACACTTAACTACTAGCAGCACTCTTTGCAGGAAACATC	978
Qy	120	SerValPheGlyValMetGluGluGlnLeuLysGlnGlyAspLysValValLeuIleSer	139
Db	979	AAAAACCTTTCTGTCAAAGAG-----CTCAAAACAGGACTTCGTGGTGTG-----CCA	1026
Qy	140	AsnHisGlnThrGluAlaAspPro-----AlaValIleAlaLeuMetLeuGluThr---	156
Db	1027	AATACCAAGTACAGGCAGACGAAACCAACAATATCGTGAACATGCTCGCTCATCTGGT	1086
Qy	157	-----ThrAsnProHisIleSerGluAsnIleIleTyrValAla	169
Db	1087	ATAGTTAACTCAAAACGCCAAGCCGGTGAAGACGCTCCAAAACGGAGCCATCTACGTAAC	1146
Qy	170	GlyAspArgVal	173
Db	1147	GGCGACGCGCATC	1158

RESULT 15
US-08-961-527-143
; Sequence 143, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

Search completed: March 16, 2004, 04:44:37
Job time : 608 secs